

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number  
**WO 01/78894 A2**

(51) International Patent Classification<sup>7</sup>: **B01L**

(21) International Application Number: PCT/US01/12245

(22) International Filing Date: 13 April 2001 (13.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/548,797 13 April 2000 (13.04.2000) US

(71) Applicant: **GENOME THERAPEUTICS CORPORATION** [US/US]; 100 Beaver Street, Waltham, MA 02154 (US).

(72) Inventor: **KEITH, Tim**; 324 North Road, Bedford, MA 01730 (US).

(74) Agents: **KOMSON, Richard, C.** et al.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

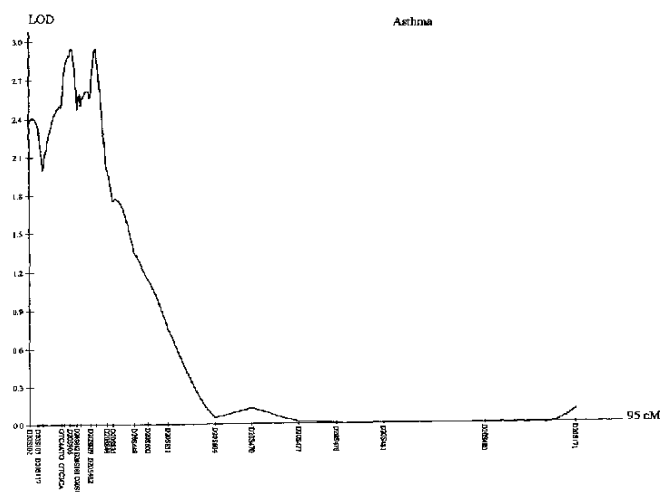
**Published:**

— without international search report and to be republished upon receipt of that report

— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES, OBESITY, AND INFLAMMATORY BOWEL DISEASE



(57) Abstract: This invention relates to genes identified from human chromosome 20p13-p12, which are associated with various diseases, including asthma. The invention also relates to the nucleotide sequences of these genes, isolated nucleic acids comprising these nucleotide sequences, and isolated polypeptides or peptides encoded thereby. The invention further relates to vectors and host cells comprising the disclosed nucleotide sequences, or fragments thereof, as well as antibodies that bind to the encoded polypeptides or peptides. Also related are ligands that modulate the activity of the disclosed genes or gene products. In addition, the invention relates to methods and compositions employing the disclosed nucleic acids, polypeptides or peptides, antibodies, and/or ligands for use in diagnostics and therapeutics for asthma and other diseases.

WO 01/78894 A2

**NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
OBESITY, AND INFLAMMATORY BOWEL DISEASE**

5

**FIELD OF THE INVENTION**

10           This invention relates to genes identified from human chromosome 20p13-p12, including Gene 216, which are associated with asthma, obesity, inflammatory bowel disease, and other human diseases. The invention also relates to the nucleotide sequences of these genes, including genomic DNA sequences, cDNA sequences, and single nucleotide polymorphisms. The  
15   invention further relates to isolated nucleic acids comprising these nucleotide sequences, and isolated polypeptides or peptides encoded thereby. Also related are expression vectors and host cells comprising the disclosed nucleic acids or fragments thereof, as well as antibodies that bind to the encoded polypeptides or peptides. The present invention further relates to ligands that  
20   modulate the activity of the disclosed genes or gene products. In addition, the invention relates to diagnostics and therapeutics for various diseases, including asthma, utilizing the disclosed nucleic acids, polypeptides or peptides, antibodies, and/or ligands.

**BACKGROUND**

25           Mouse chromosome 2 has been linked to a variety of disorders including airway hyperresponsiveness and obesity (DeSanctis et al., 1995, *Nature Genetics*, 11:150-154; Nagle et al., 1999, *Nature*, 398:148-152). This region of the mouse genome is homologous to portions of human chromosome 20 including 20p13-p12. Although human chromosome 20p13-12p has been  
30   linked to a variety of genetic disorders including diabetes insipidus, neurohypophyseal, congenital endothelial dystrophy of cornea, insomnia,

neurodegeneration with brain iron accumulation 1 (Hallervorden-Spatz syndrome), fibrodysplasia ossificans progressiva, alagille syndrome, hydrometrocolpos (McKusick-Kaufman syndrome), Creutzfeldt-Jakob disease and Gerstmann-Straussler disease (see NCBI; National Center for  
5 Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) the genes affecting these disorders have yet to be discovered. There is a need in the art for identifying specific genes relating to these disorders, as well as genes associated with obesity, lung disease, particularly, inflammatory lung disease  
10 phenotypes such as Chronic Obstructive Lung Disease (COPD), Adult Respiratory Distress Syndrome (ARDS), and asthma. Identification and characterization of such genes will make possible the development of effective diagnostics and therapeutic means to treat lung-related disorders.

#### **SUMMARY OF THE INVENTION**

15 This invention relates to Gene 216 located on human chromosome 20p13-p12. In specific embodiments, the invention relates to isolated nucleic acids comprising Gene 216 genomic sequences (e.g., SEQ ID NO:5 and SEQ ID NO:6), cDNA sequences (e.g., SEQ ID NO:1 and SEQ ID NO:3), complementary sequences, sequence variants, or fragments thereof, as  
20 described herein. The present invention also encompasses nucleic acid probes or primers useful for assaying a biological sample for the presence or expression of Gene 216. The invention further encompasses nucleic acids variants comprising single nucleotide polymorphisms (SNPs) identified in several genes, including Gene 216 (e.g., SEQ ID NO:241-288). Such SNPs  
25 can be used to diagnose diseases such as asthma, or to determine a genetic predisposition thereto. In addition, the present invention encompasses nucleic acids comprising alternate splicing variants (e.g., SEQ ID NO:2 and SEQ ID NO:350-362).

This invention also relates to vectors and host cells comprising vectors  
30 comprising the Gene 216 nucleic acid sequences disclosed herein. Such vectors can be used for nucleic acid preparations, including antisense nucleic

acids, and for the expression of encoded polypeptides or peptides. Host cells can be prokaryotic or eukaryotic cells. In specific embodiments, an expression vector comprises a DNA sequence encoding the Gene 216 polypeptide sequence (e.g., SEQ ID NO:4 or SEQ ID NO:363), sequence variants, or  
5 fragments thereof, as described herein.

The present invention further relates to isolated Gene 216 polypeptides and peptides. In specific embodiments, the polypeptides or peptides comprise the amino acid sequence of the Gene 216 (e.g., SEQ ID NO:4 or SEQ ID NO:363), sequence variants, or portions thereof, as described herein. In  
10 addition, this invention encompasses isolated fusion proteins comprising Gene 216 polypeptides or peptides.

The present invention also relates to isolated antibodies, including monoclonal and polyclonal antibodies, and antibody fragments, that are specifically reactive with the Gene 216 polypeptides, fusion proteins, or  
15 variants, or portions thereof, as disclosed herein. In specific embodiments, monoclonal antibodies are prepared to be specifically reactive with the Gene 216 polypeptide (e.g., SEQ ID NO:4 or SEQ ID NO:363) or peptides, or sequence variants thereof.

In addition, the present invention relates to methods of obtaining Gene  
20 216 polynucleotides and polypeptides, variant sequences, or fragments thereof, as disclosed herein. Also related are methods of obtaining anti-Gene 216 antibodies and antibody fragments. The present invention also encompasses methods of obtaining Gene 216 ligands, e.g., agonists, antagonists, inhibitors, and binding factors. Such ligands can be used as  
25 therapeutics for asthma and related diseases.

The present invention also relates to diagnostic methods and kits utilizing Gene 216 (wild-type, mutant, or variant) nucleic acids, polypeptides, antibodies, or functional fragments thereof. Such factors can be used, for example, in diagnostic methods and kits for measuring expression levels of  
30 Gene 216, and to screen for various Gene 216-related diseases, especially asthma. In addition, the nucleic acids described herein can be used to identify



chromosomal abnormalities affecting Gene 216, and to identify allelic variants or mutations of Gene 216 in an individual or population.

The present invention further relates to methods and therapeutics for the treatment of various diseases, including asthma. In various embodiments, therapeutics comprising the disclosed Gene 216 nucleic acids, polypeptides, antibodies, ligands, or variants, derivatives, or portions thereof, are administered to a subject to treat, prevent, or ameliorate asthma. Specifically related are therapeutics comprising Gene 216 antisense nucleic acids, monoclonal antibodies, metalloprotease inhibitors, and gene therapy vectors. Such therapeutics can be administered alone, or in combination with one or more asthma treatments.

In addition, this invention relates to non-human transgenic animals and cell lines comprising one or more of the disclosed Gene 216 nucleic acids, which can be used for drug screening, protein production, and other purposes. Also related are non-human knock-out animals and cell lines, wherein one or more endogenous Gene 216 genes (i.e., orthologs), or portions thereof, are deleted or replaced by marker genes.

This invention further relates to methods of identifying proteins that are candidates for being involved in asthma (i.e., a "candidate protein"). Such proteins are identified by a method comprising: 1) identifying a protein in a first individual having the asthma phenotype; 2) identifying a protein in a second individual not having the asthma phenotype; and 3) comparing the protein of the first individual to the protein of the second individual, wherein a) the protein that is present in the second individual but not the first individual is the candidate protein; or b) the protein that is present in a higher amount in the second individual than in the first individual is the candidate protein; or c) the protein that is present in a lower amount in the second individual than in the first individual is the candidate protein.

#### **BRIEF DESCRIPTION OF THE FIGURES**

**Figure 1** depicts the LOD Plot of Linkage to Asthma.

**Figure 2** depicts the LOD Plot of Linkage to BHR (PC20  $\leq$  4 mg/ml) &

Asthma.

**Figure 3** depicts the LOD Plot of Linkage to BHR (PC20  $\leq$  16 mg/ml) & Asthma

**Figure 4** depicts the LOD Plot of Linkage to High Total IgE & Asthma

5 **Figure 5** depicts the LOD Plot of Linkage to High Specific IgE & Asthma

**Figure 6** depicts the BAC/STS content contig map of human chromosome 20p13-p12.

**Figure 7** depicts the BAC1098L22 nucleotide sequence (SEQ ID NO:5).

10 **Figure 8** depicts the locations of single nucleotide polymorphisms, corresponding amino acid changes, and domains in the Gene 216 transcript. The exons of the transcript are marked from A to T and the size of each one is indicated. Above the exons, the 8 domains are labeled and a black bar represents the approximate location of each one. Underneath the black bars are the approximate location of the amino acid changes that have been  
15 identified. The amino acids boxed in white are the alleles that are most frequently observed. The nucleotides boxed in gray are the alleles that are most frequently observed. Single nucleotide polymorphisms are unboxed, and the polymorphism names appear underneath. The uterus cDNA clone does not contain all of Exon A, and does not contain the sequence CAG between  
20 Exon S and T.

**Figure 9** depicts alternate splice variants of Gene 216 obtained from lung tissue, including rt672 (SEQ ID NO:350), rt690 (SEQ ID NO:351), rt709 (SEQ ID NO:352), rt711 (SEQ ID NO:353), rt713 (SEQ ID NO:354), and rt720 (SEQ ID NO:355).

25 **Figure 10** depicts alternate splice variants of Gene 216 obtained from lung tissue, including rt725 (SEQ ID NO:356), rt727 (SEQ ID NO:357), rt733 (SEQ ID NO:358), rt735 (SEQ ID NO:359), rt764 (SEQ ID NO:360), rt772 (SEQ ID NO:361), and rt774 (SEQ ID NO:362).

**Figure 11** depicts the structure of the genomic sequence of Gene 216.

30 **Figure 12** depicts the alternate AG splice sequences at the junction of Intron ST and Exon T in Gene 216.

**Figure 13** depicts the promoter region of Gene 216. The Gene 216 promoter sequence is shown in SEQ ID NO:8; the Gene 216 enhancer sequence is shown in SEQ ID NO:7.

**Figure 14** depicts a dendrogram of the ADAM family members and the relationship of Gene 216 to ADAMs that possesses an active metalloprotease domain.

**Figures 15A-15C** depict Northern Blots illustrating Gene 216 expression patterns. Figures 15A-15B show Gene 216 expression in various tissue types. Figure 15C shows Gene 216 expression in bronchial smooth muscle tissue.

**Figure 16** depicts a Dot Blot that shows Gene 216 expression in various tissue types.

**Figure 17** depicts RT-PCR analysis of Gene 216 expression in primary cells from lung tissue.

**Figure 18** depicts an amino acid sequence alignment (Pileup) of 5 ADAM family members that are closely related to Gene 216. Amino acids highlighted in black show 100% identity within the Pileup; dark gray show 80% identity; and light gray show 60% identity. The boxed amino acids represent the cysteine switch, the metalloprotease domain, and the "met-turn". The labeled arrows show the locations of the 8 domains.

**Figure 19** depicts the amino acid sequence of Gene 216 (SEQ ID NO:4). Labeled arrows above the sequence denote domain and corresponding length. Black boxes represent the signal sequence and the transmembrane domain identified by hydrophobicity plots. The underlined cysteine residue at position 133 is predicted to be involved in the cysteine switch, the dashed box represents the metalloprotease domain, and the methionine underlined twice is the "met-turn". The gray boxes represent the signaling binding sites identified in the cytoplasmic tail. The amino acid changes corresponding to single nucleotide polymorphisms are indicated in bold. The alanine deleted in the uterus cDNA clone is marked within a black triangle, and if present would have been between the glutamine and the aspartic acid.

**Figure 20** depicts the Kyte-Doolittle hydrophobicity plot for the Gene

216 amino acid sequence.

**Figures 21** depicts the genomic sequence of the mouse ortholog of Gene 216.(SEQ ID NO:364).

**Figure 22** depicts the cDNA nucleotide sequence (SEQ ID NO:364) and  
5 predicted amino acid sequence (SEQ ID NO:365) of the mouse ortholog of Gene 216.

**Figure 23** depicts an amino acid sequence alignment (Pileup) of human Gene 216 polypeptide (SEQ ID NO:4) and the mouse ortholog of Gene 216 (SEQ ID NO:366). Vertical lines indicate identical amino acid residues. Dots  
10 indicate similar amino acid residues.

**Figure 24** depicts the nucleotide sequence (SEQ ID NO:1) and encoded amino acid sequence (SEQ ID NO:4) determined from the master cDNA sequence of Gene 216. The master cDNA sequence combines the sequence information from the uterine cDNA clone and 5'RACE clone. Identified single  
15 nucleotide polymorphism positions are underlined.

**Figure 25** depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the asthma phenotype in the combined US and UK populations.

**Figure 26** depicts the results of a case control study p-value plot that  
20 shows single nucleotide polymorphism association with the asthma phenotype in the US and UK populations, separately.

**Figure 27** depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the bronchial hyper-responsiveness and asthma phenotypes in the US and UK combined  
25 population.

**Figure 28** depicts the results of a case control study p-value plot that shows single nucleotide polymorphism association with the bronchial hyper-responsiveness and asthma phenotypes in the US and UK populations, separately.

**Figure 29** depicts the genomic nucleotide sequence (SEQ ID NO:6)  
30 determined for Gene 216. Identified single nucleotide polymorphism positions

are underlined.

**Figure 30** depicts the nucleotide sequence (SEQ ID NO:3) and encoded amino acid sequence (SEQ ID NO: 363) of Gene 216 determined from the uterus cDNA clone. Identified single nucleotide polymorphism positions are underlined.

**Figure 31** depicts the nucleotide sequence (SEQ ID NO:350) and encoded amino acid sequence (SEQ ID NO:337) of Gene 216 alternate splice variant rt672.

**Figure 32** depicts the nucleotide sequence (SEQ ID NO:351) and encoded amino acid sequence (SEQ ID NO:338) of Gene 216 alternate splice variant rt690.

**Figure 33** depicts the nucleotide sequence (SEQ ID NO:352) and encoded amino acid sequence (SEQ ID NO:339) of Gene 216 alternate splice variant rt709.

**Figure 34** depicts the nucleotide sequence (SEQ ID NO:353) and encoded amino acid sequence (SEQ ID NO:340) of Gene 216 alternate splice variant rt711.

**Figure 35** depicts the nucleotide sequence (SEQ ID NO:354) and encoded amino acid sequence (SEQ ID NO:341) of Gene 216 alternate splice variant rt713.

**Figure 36** depicts the nucleotide sequence (SEQ ID NO:355) and encoded amino acid sequence (SEQ ID NO:342) of Gene 216 alternate splice variant rt720.

**Figure 37** depicts the nucleotide sequence (SEQ ID NO:356) and encoded amino acid sequence (SEQ ID NO:343) of Gene 216 alternate splice variant rt725.

**Figure 38** depicts the nucleotide sequence (SEQ ID NO:357) and encoded amino acid sequence (SEQ ID NO:344) of Gene 216 alternate splice variant rt727.

**Figure 39** depicts the nucleotide sequence (SEQ ID NO:358) and encoded amino acid sequence (SEQ ID NO:345) of Gene 216 alternate splice

variant rt733.

**Figure 40** depicts the nucleotide sequence (SEQ ID NO:359) and encoded amino acid sequence (SEQ ID NO:346) of Gene 216 alternate splice variant rt735.

5       **Figure 41** depicts the nucleotide sequence (SEQ ID NO:360) and encoded amino acid sequence (SEQ ID NO:347) of Gene 216 alternate splice variant rt764.

10       **Figure 42** depicts the nucleotide sequence (SEQ ID NO:361) and encoded amino acid sequence (SEQ ID NO:348) of Gene 216 alternate splice variant rt772.

**Figure 43** depicts the nucleotide sequence (SEQ ID NO:362) and encoded amino acid sequence (SEQ ID NO:349) of Gene 216 alternate splice variant rt774.

#### **DETAILED DESCRIPTION OF THE INVENTION**

15       Gene 216 was identified by extensive analysis of the region of human chromosome 20p13-p12 associated with airway hyperresponsiveness, asthma, and atopy. This region has also been implicated in other diseases such as obesity (Wilson, 1999, *Arch. Intern. Med.* **159**:2513-4). Bronchial asthma, furthermore, has been linked to intestinal conditions such as inflammatory  
20       bowel disease (B. Wallaert et al., 1995, *J. Exp. Med.* **182**:1897-1904). Thus, there was a need to identify and isolate the gene(s) associated with this region of human chromosome 20.

#### **Definitions**

25       To aid in the understanding of the specification and claims, the following definitions are provided.

      "Disorder region" refers to a portion of the human chromosome 20 bounded by the markers D20S502 and D20S851. A "disorder-associated" nucleic acid or polypeptide sequence refers to a nucleic acid sequence that maps to region 20p13-p12 or the polypeptides encoded therein (e.g., Gene  
30       216 nucleic acids, and polypeptides). For nucleic acids, this encompasses sequences that are identical or complementary to the Gene 216 sequence, as

well as sequence-conservative, function-conservative, and non-conservative variants thereof. For polypeptides, this encompasses sequences that are identical to the Gene 216 polypeptide, as well as function-conservative and non-conservative variants thereof. Included are naturally-occurring mutations  
5 of Gene 216 causative of respiratory diseases or obesity, such as but not limited to mutations which cause altered protein levels or stability (e.g., decreased levels, increased levels, expression in an inappropriate tissue type, increased stability, and decreased stability).

As used herein, the "reference sequence" for Gene 216 is BAC1098L22  
10 (SEQ ID NO:5). The BAC1098L22 sequence is also the source of the disclosed Gene 216 genomic sequence (SEQ ID NO:6). "Variant" sequences refer to nucleotide sequences (and the encoded amino acid sequences) that differ from the reference sequence at one or more positions. Non-limiting examples of variant sequences include the disclosed Gene 216 single  
15 nucleotide polymorphisms (SNPs), alternate splice variants, and the amino acid sequences encoded by these variants.

"Sequence-conservative" variants are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position (i.e., silent mutations). "Function-  
20 conservative" variants are those in which a change in one or more nucleotides in a given codon position results in a polypeptide sequence in which a given amino acid residue in the polypeptide has been replaced by a conservative amino acid substitution as described in detail herein. "Function-conservative" variants also include analogs of a given polypeptide and any polypeptides that  
25 have the ability to elicit antibodies specific to a designated polypeptide. "Non-conservative" variants are those in which a change in one or more nucleotides in a given codon position results in a polypeptide sequence in which a given amino acid residue in a polypeptide has been replaced by a non-conservative amino acid substitution as described hereinbelow. "Non-conservative" variants  
30 also include polypeptides comprising non-conservative amino acid substitutions.

As used herein, the term "ortholog" denotes a gene or polypeptide obtained from one species that has homology to an analogous gene or polypeptide from a different species. The term "paralog" denotes a gene or polypeptide obtained from a given species that has homology to a distinct gene or polypeptide from that same species. For example, the disclosed mouse and human Gene 216 sequences are orthologs, whereas human Gene 216 and human ADAM 19 are paralogs.

"Nucleic acid or "polynucleotide" as used herein refers to purine- and pyrimidine-containing polymers of any length, either polyribonucleotides or polydeoxyribonucleotide or mixed polyribo-polydeoxyribonucleotides. This includes single-and double-stranded molecules, i.e., DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases.

As used herein, "isolated" nucleic acids are nucleic acids separated away from other components (e.g., DNA, RNA, and protein) with which they are associated (e.g., as obtained from cells, chemical synthesis systems, or phage or nucleic acid libraries). Isolated nucleic acids are at least 60% free, preferably 75% free, and most preferably 90% free from other associated components. In accordance with the present invention, isolated nucleic acids can be obtained by methods described herein, or other established methods, including isolation from natural sources (e.g., cells, tissues, or organs), chemical synthesis, recombinant methods, combinations of recombinant and chemical methods, and library screening methods.

Nucleic acids referred to herein as "recombinant" are nucleic acids which have been produced by recombinant DNA methodology, including those nucleic acids that are generated by procedures which rely upon a method of artificial replication, such as the polymerase chain reaction (PCR) and/or cloning into a vector using restriction enzymes. Portions of recombinant nucleic acids which code for polypeptides can be identified and isolated by, for example, the method of M. Jasin et al., U.S. Patent No. 4,952,501.



A "coding sequence" or a "protein-coding sequence" is a polynucleotide sequence capable of being transcribed into mRNA and/or capable of being translated into a polypeptide or peptide. The boundaries of the coding sequence are typically determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus.

A "complement" of a nucleic acid sequence as used herein refers to the "antisense" sequence that participates in Watson-Crick base-pairing with the original sequence.

A "probe" or "primer" refers to a nucleic acid or oligonucleotide that forms a hybrid structure with a sequence in a target region due to complementarity of the probe or primer sequence to at least one portion of the target region sequence.

Nucleic acids are "hybridizable" to each other when at least one strand of the nucleic acid can anneal to another nucleic acid strand under defined stringency conditions. Hybridization requires that the two nucleic acids contain substantially complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementarity, and can be determined in accordance with the methods described herein.

As used herein, "portion" and "fragment" are synonymous. A "portion" as used with regard to a nucleic acid or polynucleotide, refers to fragments of that nucleic acid or polynucleotide. The fragments can range in size from 8 nucleotides to all but one nucleotide of the entire Gene 216 sequence. Preferably, The fragments are at least 8 to 10 nucleotides in length; more preferably at least 12 nucleotides in length; still more preferably at least 15 to 20 nucleotides in length; yet more preferably at least 25 nucleotides in length; and most preferably at least 35 to 55 nucleotides in length.

"cDNA" refers to complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus, a "cDNA clone" means a duplex DNA sequence

complementary to an RNA molecule of interest, included in a cloning vector or PCR amplified. This term includes genes from which the intervening sequences have been removed.

"Cloning" refers to the use of recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired gene, it is necessary to use methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient host cells.

"cDNA library" refers to a collection of recombinant DNA molecules containing cDNA inserts that together comprise essentially all of the expressed genes of an organism. A cDNA library can be prepared by methods known to one skilled in the art (see, e.g., Cowell and Austin, 1997, "cDNA Library Protocols," *Methods in Molecular Biology*). Generally, RNA is first isolated from the cells of the desired organism, and the RNA is used to prepare cDNA molecules.

"Cloning vector" refers to a plasmid or phage DNA or other DNA that is able to replicate in a host cell. The cloning vector is typically characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the DNA, which may contain a marker suitable for use in the identification of cells containing the vector.

"Regulatory sequence" refers to a nucleic acid sequence that controls or regulates expression of structural genes when operably linked to those genes. These include, for example, the lac system, the trp system, major operator and promoter regions of the phage lambda, the control region of fd coat protein and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells. Regulatory sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host, and may contain transcriptional elements such

as enhancer elements, termination sequences, tissue-specificity elements and/or translational initiation and termination sites.

"Expression vector" refers to a vehicle or plasmid that is capable of expressing a gene that has been cloned into it, after transformation or  
5 integration in a host cell. The cloned gene is usually placed under the control of (i.e., operably linked to) a regulatory sequence.

"Operably linked" means that the promoter controls the initiation of expression of the gene. A promoter is operably linked to a sequence of proximal DNA if upon introduction into a host cell the promoter determines the  
10 transcription of the proximal DNA sequence(s) into one or more species of RNA. A promoter is operably linked to a DNA sequence if the promoter is capable of initiating transcription of that DNA sequence.

"Host" includes prokaryotes and eukaryotes. The term includes an organism or cell that is the recipient of an expression vector (e.g.,  
15 autonomously replicating or integrating vector).

"Amplification" of nucleic acids refers to methods such as polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods based on the use of Q-beta replicase. These methods are well known in the art and described, for example, in U.S. Patent Nos.  
20 4,683,195 and 4,683,202. Reagents and hardware for conducting PCR are commercially available. Primers useful for amplifying sequences from the disorder region are preferably complementary to, and preferably hybridize specifically to, sequences in the 20p13-p12 region or in regions that flank a target region therein. Gene 216 generated by amplification may be sequenced  
25 directly. Alternatively, the amplified sequence(s) may be cloned prior to sequence analysis.

"Gene" refers to a DNA sequence that encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide, polypeptide, or protein. The term "gene" as used herein with reference to  
30 genomic DNA includes intervening, non-coding regions, as well as regulatory regions, and can include 5' and 3' ends.

A gene sequence is "wild-type" if such sequence is usually found in individuals unaffected by the disease or condition of interest. However, environmental factors and other genes can also play an important role in the ultimate determination of the disease. In the context of complex diseases involving multiple genes ("oligogenic disease"), the "wild type", or normal sequence can also be associated with a measurable risk or susceptibility, receiving its reference status based on its frequency in the general population. As used herein, "wild-type Gene 216" refers to the reference sequence, BAC1098L22 (SEQ ID NO:5). The wild-type Gene 216 sequence was used to identify the variants (single nucleotide polymorphisms) described in detail herein.

A gene sequence is a "mutant" sequence if it differs from the wild-type sequence. For example, a Gene 216 nucleic acid containing a single nucleotide polymorphism is a mutant sequence. In some cases, the individual carrying such gene has increased susceptibility toward the disease or condition of interest. In other cases, the "mutant" sequence might also refer to a sequence that decreases the susceptibility toward a disease or condition of interest, and thus acting in a protective manner. Also a gene is a "mutant" gene if too much ("overexpressed") or too little ("underexpressed") of such gene is expressed in the tissues in which such gene is normally expressed, thereby causing the disease or condition of interest.

A nucleic acid or fragment thereof is "substantially homologous" to another if, when optimally aligned (with appropriate nucleotide insertions and/or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least 60% of the nucleotide bases, usually at least 70%, more usually at least 80%, preferably at least 90%, and more preferably at least 95-98% of the nucleotide bases.

Alternatively, substantial homology exists when a nucleic acid or fragment thereof will hybridize, under selective hybridization conditions, to another nucleic acid (or a complementary strand thereof). Selectivity of hybridization exists when hybridization which is substantially more selective

than total lack of specificity occurs. Typically, selective hybridization will occur when there is at least about 55% sequence identity over a stretch of at least about nine or more nucleotides, preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90% (M. Kanehisa,  
5 1984, *Nucl. Acids Res.* **11**:203-213). The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will often be over a stretch of at least 14 nucleotides, usually at least 20 nucleotides, more usually at least 24 nucleotides, typically at least 28 nucleotides, more typically at least 32 nucleotides, and preferably at least 36 or more nucleotides.

10 As used herein, the terms "protein" and "polypeptide" are synonymous. "Peptides" are defined as fragments or portions of polypeptides, preferably fragments or portions having at least one functional activity (e.g., proteolysis, adhesion, fusion, antigenic, or intracellular activity) as the complete polypeptide sequence.

15 "Isolated" polypeptides or peptides are those that are separated from other components (e.g., DNA, RNA, and other polypeptides or peptides) with which they are associated (e.g., as obtained from cells, translation systems, or chemical synthesis systems). In a preferred embodiment, isolated polypeptides or peptides are at least 10% pure; more preferably, 80 or 90%  
20 pure. Isolated polypeptides and peptides include those obtained by methods described herein, or other established methods, including isolation from natural sources (e.g., cells, tissues, or organs), chemical synthesis, recombinant methods, or combinations of recombinant and chemical methods. Proteins or polypeptides referred to herein as "recombinant" are proteins or polypeptides  
25 produced by the expression of recombinant nucleic acids.

A "portion" as used herein with regard to a protein or polypeptide, refers to fragments of that protein or polypeptide. The fragments can range in size from 5 amino acid residues to all but one residue of the entire protein sequence. Thus, a portion or fragment can be at least 5, 5-50, 50-100, 100-  
30 200, 200-400, 400-800, or more consecutive amino acid residues of a Gene 216 protein or polypeptide, for example, SEQ ID NO:4 or SEQ ID NO:363.

An "immunogenic component", is a moiety that is capable of eliciting a humoral and/or cellular immune response in a host animal.

An "antigenic component" is a moiety that binds to its specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

5       A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including, without limitation, plasma, serum, cerebrospinal fluid, lymph, tears, saliva, milk, pus, and tissue exudates and secretions) or from *in vitro* cell culture constituents, as well as samples obtained from, for example, a laboratory procedure.

10       "Antibodies" refer to polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, that can bind to asthma proteins and fragments thereof or to nucleic acid sequences from the 20p13-p12 region, particularly from the asthma locus or a portion thereof.

15       The term antibody is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Proteins may be prepared synthetically in a protein synthesizer and coupled to a carrier molecule and injected over several months into rabbits. Rabbit sera is tested for immunoreactivity to the protein or fragment. Monoclonal antibodies may be made by injecting mice with the  
20       proteins, or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with protein or fragments thereof. (Harlow et al., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). These antibodies will be useful in assays as well as therapeutics.

25       "Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings  
30       of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (A.M. Lesk (ed),

1988, *Computational Molecular Biology*, Oxford University Press, NY; D.W. Smith (ed), 1993, *Biocomputing. Informatics and Genome Projects*, Academic Press, NY; A.M. Griffin and H.G. Griffin, H. G (eds), 1994, *Computer Analysis of Sequence Data*, Part I, Humana Press, NJ; G. von Heinje, 1987, *Sequence*  
5 *Analysis in Molecular Biology*, Academic Press; and M. Gribskov and J. Devereux (eds), 1991, *Sequence Analysis Primer*, M Stockton Press, NY; H. Carillo and D. Lipman, 1988, *SIAM J. Applied Math.*, **48**:1073.

Technical and scientific terms used herein have the meanings commonly understood by one of ordinary skill in the art to which the present  
10 invention pertains, unless otherwise defined. Reference is made herein to various methodologies known to those of skill in the art. Publications and other materials setting forth such known methodologies to which reference is made are incorporated herein by reference in their entireties as though set forth in full.

15 Standard reference works setting forth the general principles of recombinant DNA technology include J. Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; P.B. Kaufman et al., (eds), 1995, *Handbook of Molecular and Cellular Methods in Biology and Medicine*, CRC Press, Boca  
20 Raton; M.J. McPherson (ed), 1991, *Directed Mutagenesis: A Practical Approach*, IRL Press, Oxford; J. Jones, 1992, *Amino Acid and Peptide Synthesis*, Oxford Science Publications, Oxford; B.M. Austen and O.M.R. Westwood, 1991, *Protein Targeting and Secretion*, IRL Press, Oxford; D.N Glover (ed), 1985, *DNA Cloning*, Volumes I and II; M.J. Gait (ed), 1984,  
25 *Oligonucleotide Synthesis*; B.D. Hames and S.J. Higgins (eds), 1984, *Nucleic Acid Hybridization*; Wu and Grossman (eds), *Methods in Enzymology* (Academic Press, Inc.), Vol. 154 and Vol. 155; Quirke and Taylor (eds), 1991, *PCR-A Practical Approach*; Hames and Higgins (eds), 1984, *Transcription and Translation*; R.I. Freshney (ed), 1986, *Animal Cell Culture; Immobilized Cells*  
30 *and Enzymes*, 1986, IRL Press; Perbal, 1984, *A Practical Guide to Molecular Cloning*; J. H. Miller and M. P. Calos (eds), 1987, *Gene Transfer Vectors for*

*Mammalian Cells*, Cold Spring Harbor Laboratory Press; M.J. Bishop (ed), 1998, *Guide to Human Genome Computing*, 2d Ed., Academic Press, San Diego, CA; L.F. Peruski and A.H. Peruski, 1997, *The Internet and the New Biology: Tools for Genomic and Molecular Research*, American Society for  
5 Microbiology, Washington, D.C.

Standard reference works setting forth the general principles of immunology include S. Sell, 1996, *Immunology, Immunopathology & Immunity*, 5th Ed., Appleton & Lange, Publ., Stamford, CT; D. Male et al., 1996, *Advanced Immunology*, 3d Ed., Times Mirror Int'l Publishers Ltd., Publ.,  
10 London; D.P. Stites and A.I. Terr, 1991, *Basic and Clinical Immunology*, 7th Ed., Appleton & Lange, Publ., Norwalk, CT; and A.K. Abbas et al., 1991, *Cellular and Molecular Immunology*, W. B. Saunders Co., Publ., Philadelphia, PA. Any suitable materials and/or methods known to those of skill can be utilized in carrying out the present invention; however, preferred materials  
15 and/or methods are described. Materials, reagents, and the like to which reference is made in the following description and examples are generally obtainable from commercial sources, and specific vendors are cited herein.

### **Nucleic Acids**

The present invention relates to isolated Gene 216 nucleic acids  
20 comprising genomic DNA within BAC RPCI\_1098L22 (e.g., SEQ ID NO:5), the corresponding cDNA sequences (e.g., SEQ ID NO:1 or SEQ ID NO:3), RNA, fragments of the genomic, cDNA, or RNA nucleic acids comprising 20, 40, 60, 100, 200, 500 or more contiguous nucleotides, and the complements thereof. Closely related variants are also included as part of this invention, as well as  
25 nucleic acids sharing at least 50, 60, 70, 80, or 90% identity with the nucleic acids described above, and nucleic acids which would be identical to a Gene 216 nucleic acids except for one or a few substitutions, deletions, or additions.

The invention also relates to isolated nucleic acids comprising regions required for accurate expression of Gene 216 (e.g., Gene 216 promoter (e.g.,  
30 SEQ ID NO:8), enhancer (e.g., SEQ ID NO:7), and polyadenylation sequences). In a preferred embodiment, the present invention is directed to



at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:6. More particularly, embodiments of this invention include the BAC clone containing segments of Gene 216 including RPCI\_1098L22 as set forth in SEQ ID NO:5 (Figure 7).

5           The invention further relates to nucleic acids (e.g., DNA or RNA) that hybridize to a) a nucleic acid encoding a Gene 216 polypeptide, such as a nucleic acid having the sequence of SEQ ID NO:1 or SEQ ID NO:6; b) sequence-conservative, function-conservative, and non-conservative variants of (a); and c) fragments or portions of (a) or (b). Nucleic acids that hybridize  
10       to the sequence of SEQ ID NO:1 or SEQ ID NO:6 can be double- or single-stranded. Hybridization to the sequence of SEQ ID NO:1 or SEQ ID NO:6 includes hybridization to the strand shown or its complementary strand.

          The present invention also relates to nucleic acids that encode a polypeptide having the amino acid sequence of SEQ ID NO:4 or SEQ ID  
15       NO:363, or functional equivalents thereof. A functional equivalent of a Gene 216 protein includes fragments or variants that perform at least on characteristic function of the Gene 216 protein (e.g., proteolysis, adhesion, fusion, antigenic, or intracellular activity). Preferably, a functional equivalent will share at least 65% sequence identity with the Gene 216 polypeptide.

20           In preferred embodiments, nucleic acids of the present invention share at least 50%, preferably at least 60-70%, more preferably at least 70-80% sequence identity, and even more preferably at least 90-100% sequence identity with the sequences of SEQ ID NO:1 or SEQ ID NO:6, or fragments or portions thereof. Sequence identity calculations can be performed using  
25       computer programs, hybridization methods, or calculations. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package, BLASTN, BLASTX, TBLASTX, and FASTA (J. Devereux et al., 1984, *Nucleic Acids Research* **12**(1):387; S.F. Altschul et al., 1990, *J. Molec. Biol.* **215**:403-  
30       410; W. Gish and D.J. States, 1994, *Nature Genet.* **3**:266-272; W.R. Pearson and D.J. Lipman, 1988, *Proc Natl. Acad. Sci. USA* **85**(8):2444-8). The BLAST

programs are publicly available from NCBI and other sources . The well-known Smith Waterman algorithm may also be used to determine identity.

For example, nucleotide sequence identity can be determined by comparing a query sequences to sequences in publicly available sequence  
5 databases (NCBI) using the BLASTN2 algorithm (S.F. Altschul et al., 1997, *Nucl. Acids Res.*, **25**:3389-3402). The parameters for a typical search are:  $E = 0.05$ ,  $v = 50$ ,  $B = 50$ , wherein  $E$  is the expected probability score cutoff,  $V$  is the number of database entries returned in the reporting of the results, and  $B$  is the number of sequence alignments returned in the reporting of the results  
10 (S.F. Altschul et al., 1990, *J. Mol. Biol.*, **215**:403-410).

In another approach, nucleotide sequence identity can be calculated using the following equation: % identity = (number of identical nucleotides) / (alignment length in nucleotides) \* 100. For this calculation, alignment length includes internal gaps but not includes terminal gaps. Alternatively, nucleotide  
15 sequence identity can be determined experimentally using the specific hybridization conditions described below.

In accordance with the present invention, polynucleotide alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, insertion, or modification  
20 (e.g., via RNA or DNA analogs). Alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. Alterations of a polynucleotide sequence of SEQ ID NO:1 or SEQ  
25 ID NO:6 may create nonsense, missense, or frameshift mutations in this coding sequence, and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

Such altered nucleic acids, including DNA or RNA, can be detected and isolated by hybridization under high stringency conditions or moderate  
30 stringency conditions, for example, which are chosen to prevent hybridization of nucleic acids having non-complementary sequences. "Stringency

conditions" for hybridizations is a term of art which refers to the conditions of temperature and buffer concentration which permit hybridization of a particular nucleic acid to another nucleic acid in which the first nucleic acid may be perfectly complementary to the second, or the first and second may share  
5 some degree of complementarity which is less than perfect.

For example, certain high stringency conditions can be used which distinguish perfectly complementary nucleic acids from those of less complementarity. "High stringency conditions" and "moderate stringency conditions" for nucleic acid hybridizations are explained in F.M. Ausubel et al.  
10 (eds), 1995, *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., New York, NY, the teachings of which are hereby incorporated by reference.

In particular, see pages 2.10.1-2.10.16 (especially pages 2.10.8-2.10.11) and pages 6.3.1-6.3.6. The exact conditions which determine the stringency of hybridization depend not only on ionic strength, temperature and the  
15 concentration of destabilizing agents such as formamide, but also on factors such as the length of the nucleic acid sequence, base composition, percent mismatch between hybridizing sequences and the frequency of occurrence of subsets of that sequence within other non-identical sequences. Thus, high or moderate stringency conditions can be determined empirically.

By varying hybridization conditions from a level of stringency at which  
20 no hybridization occurs to a level at which hybridization is first observed, conditions which will allow a given sequence to hybridize with the most similar sequences in the sample can be determined. Preferably the hybridizing sequences will have 60-70% sequence identity, more preferably 70-85%  
25 sequence identity, and even more preferably 90-100% sequence identity.

Typically, the hybridization reaction is initially performed under conditions of low stringency, followed by washes of varying, but higher stringency. Reference to hybridization stringency, e.g., high, moderate, or low stringency, typically relates to such washing conditions. Hybridization  
30 conditions are based on the melting temperature ( $T_m$ ) of the nucleic acid probe or primer and are typically classified by degree of stringency of the conditions

under which hybridization is measured (Ausubel et al., 1995). For example, high stringency hybridization typically occurs at about 5-10% C below the  $T_m$ ; moderate stringency hybridization occurs at about 10-20% below the  $T_m$ ; and low stringency hybridization occurs at about 20-25% below the  $T_m$ . The melting temperature can be approximated by the formulas as known in the art, depending on a number of parameters, such as the length of the hybrid or probe in number of nucleotides, or hybridization buffer ingredients and conditions. As a general guide,  $T_m$  decreases approximately 1°C with every 1% decrease in sequence identity at any given SSC concentration. Generally, doubling the concentration of SSC results in an increase in  $T_m$  of ~17°C. Using these guidelines, the washing temperature can be determined empirically for moderate or low stringency, depending on the level of mismatch sought.

High stringency hybridization conditions are typically carried out at 65 to 68°C in 0.1 X SSC and 0.1% SDS. Highly stringent conditions allow hybridization of nucleic acid molecules having about 95 to 100% sequence identity. Moderate stringency hybridization conditions are typically carried out at 50 to 65°C in 1 X SSC and 0.1% SDS. Moderate stringency conditions allow hybridization of sequences having at least about 80 to 95% nucleotide sequence identity. Low stringency hybridization conditions are typically carried out at 40 to 50°C in 6 X SSC and 0.1% SDS. Low stringency hybridization conditions allow detection of specific hybridization of nucleic acid molecules having at least about 50 to 80% nucleotide sequence identity.

For example, high stringency conditions can be attained by hybridization in 50% formamide, 5 X Denhardt's solution, 5 X SSPE or SSC (1 X SSPE buffer comprises 0.15 M NaCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1 mM EDTA; 1 X SSC buffer comprises 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.2% SDS at about 42°C, followed by washing in 1 X SSPE or SSC and 0.1% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C. Moderate stringency conditions can be attained, for example, by hybridization in 50% formamide, 5 X Denhardt's solution, 5 X SSPE or SSC, and 0.2% SDS at 42°C to about 50°C, followed by washing in 0.2 X SSPE or

SSC and 0.2% SDS at a temperature of at least about 42°C, preferably about 55°C, more preferably about 65°C. Low stringency conditions can be attained, for example, by hybridization in 10% formamide, 5 X Denhardt's solution, 6 X SSPE or SSC, and 0.2% SDS at 42°C, followed by washing in 1 X SSPE or  
5 SSC, and 0.2% SDS at a temperature of about 45°C, preferably about 50°C in 4 X SSC at 60°C for 30 min.

High stringency hybridization procedures typically (1) employ low ionic strength and high temperature for washing, such as 0.015 M NaCl/ 0.0015 M sodium citrate, pH 7.0 (0.1 X SSC) with 0.1% sodium dodecyl sulfate (SDS) at  
10 50°C; (2) employ during hybridization 50% (vol/vol) formamide with 5 X Denhardt's solution (0.1% weight/volume highly purified bovine serum albumin/0.1% wt/vol Ficoll/0.1% wt/vol polyvinylpyrrolidone), 50 mM sodium phosphate buffer at pH 6.5 and 5 X SSC at 42°C; or (3) employ hybridization  
15 with 50% formamide, 5 X SSC, 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 X Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 X SSC and 0.1% SDS.

In one particular embodiment, high stringency hybridization conditions may be attained by:

20 -- Prehybridization treatment of the support (e.g. nitrocellulose filter or nylon membrane), to which is bound the nucleic acid capable of hybridizing with any of the sequences of the invention, is carried out at 65°C for 6 hr with a solution having the following composition: 4 X SSC, 10 X Denhardt's (1 X Denhardt's comprises 1% Ficoll, 1% polyvinylpyrrolidone, 1% BSA (bovine  
25 serum albumin); 1 X SSC comprises of 0.15 M of NaCl and 0.015 M of sodium citrate, pH 7);

-- Replacement of the pre-hybridization solution in contact with the support by a buffer solution having the following composition: 4 X SSC, 1 X Denhardt's, 25 mM NaPO<sub>4</sub>, pH 7, 2 mM EDTA, 0.5% SDS, 100 µg/ml of  
30 sonicated salmon sperm DNA containing a nucleic acid derived from the

sequences of the invention as probe, in particular a radioactive probe, and previously denatured by a treatment at 100°C for 3 min;

-- Incubation for 12 hr at 65°C;

-- Successive washings with the following solutions: 1) four washings with  
5 2 X SSC, 1 X Denhardt's, 0.5% SDS for 45 min at 65°C; 2) two washings with 0.2 X SSC, 0.1 X SSC for 45 min at 65°C; and 3) 0.1 x SSC, 0.1% SDS for 45 min at 65°C.

Additional examples of high, medium, and low stringency conditions can be found in Sambrook et al., 1989. Exemplary conditions are also described  
10 in M.H. Krause and S.A. Aaronson, 1991, *Methods in Enzymology*, **200**:546-556; Ausubel et al., 1995. It is to be understood that the low, moderate and high stringency hybridization/washing conditions may be varied using a variety of ingredients, buffers, and temperatures well known to and practiced by the skilled practitioner.

15 Isolated nucleic acids that are characterized by their ability to hybridize to (a) a nucleic acid encoding a Gene 216 polypeptide, such as the nucleic acids depicted as SEQ ID NO:1 or SEQ ID NO:6, b) the complement of (a), (c) or a portion of (a) or (b) (e.g., under high or moderate stringency conditions), may further encode a protein or polypeptide having at least one function  
20 characteristic of a Gene 216 polypeptide, such as proteolysis, adhesion, fusion, and intracellular activity, or binding of antibodies that also bind to non-recombinant Gene 216 protein or polypeptide. The catalytic or binding function of a protein or polypeptide encoded by the hybridizing nucleic acid may be detected by standard enzymatic assays for activity or binding (e.g., assays that  
25 measure the binding of a transit peptide or a precursor, or other components of the translocation machinery). Enzymatic assays, complementation tests, or other suitable methods can also be used in procedures for the identification and/or isolation of nucleic acids which encode a polypeptide having the amino acid sequence of SEQ ID NO:4 or SEQ ID NO:363, or a functional equivalent  
30 of this polypeptide. The antigenic properties of proteins or polypeptides encoded by hybridizing nucleic acids can be determined by immunological

methods employing antibodies that bind to a Gene 216 polypeptide such as immunoblot, immunoprecipitation and radioimmunoassay. PCR methodology, including RAGE (Rapid Amplification of Genomic DNA Ends), can also be used to screen for and detect the presence of nucleic acids which encode Gene  
5 216-like proteins and polypeptides, and to assist in cloning such nucleic acids from genomic DNA. PCR methods for these purposes can be found in M.A. Innis et al., 1990, *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA., incorporated herein by reference.

It is understood that, as a result of the degeneracy of the genetic code,  
10 many nucleic acid sequences are possible which encode a Gene 216-like protein or polypeptide. Some of these will share little identity to the nucleotide sequences of any known or naturally-occurring Gene 216-like gene but can be used to produce the proteins and polypeptides of this invention by selection of combinations of nucleotide triplets based on codon choices. Such variants,  
15 while not hybridizable to a naturally-occurring Gene 216 gene under conditions of high stringency, are contemplated within this invention.

Also encompassed by the present invention are alternate splice variants produced by differential processing of the primary transcript(s) from Gene 216 genomic DNA. An alternate splice variant may comprise, for example, the  
20 sequence of any one of SEQ ID NO:2 and SEQ ID NO:350-362. Alternate splice variants can also comprise other combinations of introns/exons of SEQ ID NO:1 or SEQ ID NO:6, which can be determined by those of skill in the art. Alternate splice variants can be determined experimentally, for example, by isolating and analyzing cellular RNAs (e.g., Southern blotting or PCR), or by  
25 screening cDNA libraries using the Gene 216 nucleic acid probes or primers described herein. In another approach, alternate splice variants can be predicted using various methods, computer programs, or computer systems available to practitioners in the field.

General methods for splice site prediction can be found in Nakata, 1985,  
30 *Nucleic Acids Res.* **13**:5327-5340. In addition, splice sites can be predicted using, for example, the GRAIL™ (E.C. Uberbacher and R.J. Mural, 1991, *Proc.*

*Natl. Acad. Sci. USA*, **88**:11261-11265; E.C. Uberbacher, 1995, *Trends Biotech.*, **13**:497-500; <http://grail.lsd.ornl.gov/grailexp>); GenView (L. Milanesi et al., 1993, *Proceedings of the Second International Conference on Bioinformatics, Supercomputing, and Complex Genome Analysis*, H.A. Lim et al. (eds), World Scientific Publishing, Singapore, pp. 573-588; [http://l25.itba.mi.cnr.it/~webgene/wwwgene\\_help.html](http://l25.itba.mi.cnr.it/~webgene/wwwgene_help.html)); SpliceView (<http://www.itba.mi.cnr.it/webgene>); and HSPL (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163; V.V. Solovyev et al., 1994, "The Prediction of Human Exons by Oligonucleotide Composition and Discriminant Analysis of Spliceable Open Reading Frames," R. Altman et al. (eds), *The Second International conference on Intelligent systems for Molecular Biology*, AAAI Press, Menlo Park, CA, pp. 354-362; V.V. Solovyev et al., 1993, "Identification Of Human Gene Functional Regions Based On Oligonucleotide Composition," L. Hunter et al. (eds), *In Proceedings of First International conference on Intelligent System for Molecular Biology*, Bethesda, pp. 371-379) computer systems.

Additionally, computer programs such as GeneParser (E.E. Snyder and G.D. Stormo, 1995, *J. Mol. Biol.* **248**: 1-18; E.E. Snyder and G.D. Stormo, 1993, *Nucl. Acids Res.* **21**(3): 607-613; <http://mcdb.colorado.edu/~eesnyder/GeneParser.html>); MZEF (M.Q. Zhang, 1997, *Proc. Natl. Acad. Sci. USA*, **94**:565-568; <http://argon.cshl.org/genefinder>); MORGAN (S. Salzberg et al., 1998, *J. Comp. Biol.* **5**:667-680; S. Salzberg et al. (eds), 1998, *Computational Methods in Molecular Biology*, Elsevier Science, New York, NY, pp. 187-203); VEIL (J. Henderson et al., 1997, *J. Comp. Biol.* **4**:127-141); GeneScan (S. Tiwari et al., 1997, *CABIOS (Bioinformatics)* **13**: 263-270); GeneBuilder (L. Milanesi et al., 1999, *Bioinformatics* **15**:612-621); Eukaryotic GeneMark (J. Besemer et al., 1999, *Nucl. Acids Res.* **27**:3911-3920); and FEXH (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163). In addition, splice sites (i.e., former or potential splice sites) in cDNA sequences can be predicted using, for example, the RNASPL (V.V. Solovyev et al., 1994, *Nucleic Acids Res.* **22**:5156-5163); or INTRON (A. Globek et al., 1991, INTRON version 1.1 manual, Laboratory of Biochemical Genetics, NIMH, Washington, D.C.)



programs.

The present invention also encompasses naturally-occurring polymorphisms of Gene 216. As will be understood by those in the art, the genomes of all organisms undergo spontaneous mutation in the course of their continuing evolution generating variant forms of gene sequences (Gusella, 1986, *Ann. Rev. Biochem.* **55**:831-854). Restriction fragment length polymorphisms (RFLPs) include variations in DNA sequences that alter the length of a restriction fragment in the sequence (Botstein et al., 1980, *Am. J. Hum. Genet.* **32**, 314-331 (1980). RFLPs have been widely used in human and animal genetic analyses (see WO 90/13668; WO90/11369; Donis-Keller, 1987, *Cell* **51**:319-337; Lander et al., 1989, *Genetics* **121**: 85-99). Short tandem repeats (STRs) include tandem di-, tri- and tetranucleotide repeated motifs, also termed variable number tandem repeat (VNTR) polymorphisms. VNTRs have been used in identity and paternity analysis (U.S. Pat. No. 5,075,217; Armour et al., 1992, *FEBS Lett.* **307**:113-115; Horn et al., WO 91/14003; Jeffreys, EP 370,719), and in a large number of genetic mapping studies.

Single nucleotide polymorphisms (SNPs) are far more frequent than RFLPs, STRs, and VNTRs. SNPs may occur in protein coding (e.g., exon), or non-coding (e.g., intron, 5'UTR, 3'UTR) sequences. SNPs in protein coding regions may comprise silent mutations that do not alter the amino acid sequence of a protein. Alternatively, SNPs in protein coding regions may produce conservative or non-conservative amino acid changes, described in detail below. In some cases, SNPs may give rise to the expression of a defective or other variant protein and, potentially, a genetic disease. SNPs within protein-coding sequences can give rise to genetic diseases, for example, in the  $\beta$ -globin (sickle cell anemia) and CFTR (cystic fibrosis) genes. In non-coding sequences, SNPs may also result in defective protein expression (e.g., as a result of defective splicing). Other single nucleotide polymorphisms have no phenotypic effects.

Single nucleotide polymorphisms can be used in the same manner as

RFLPs and VNTRs, but offer several advantages. Single nucleotide polymorphisms tend to occur with greater frequency and are typically spaced more uniformly throughout the genome than other polymorphisms. Also, different SNPs are often easier to distinguish than other types of polymorphisms (e.g., by use of assays employing allele-specific hybridization probes or primers). In one embodiment of the present invention, a Gene 216 nucleic acid contains at least one SNP as set forth in Table 10, herein below. Various combinations of these SNPs are also encompassed by the invention. In a preferred aspect, a Gene 216 SNP is associated with a lung-related disorder, such as asthma.

The nucleic acid sequences of the present invention may be derived from a variety of sources including DNA, cDNA, synthetic DNA, synthetic RNA, or combinations thereof. Such sequences may comprise genomic DNA, which may or may not include naturally occurring introns. Moreover, such genomic DNA may be obtained in association with promoter regions or poly (A) sequences. The sequences, genomic DNA, or cDNA may be obtained in any of several ways. Genomic DNA can be extracted and purified from suitable cells by means well known in the art. Alternatively, mRNA can be isolated from a cell and used to produce cDNA by reverse transcription or other means.

The nucleic acids described herein are used in the methods of the present invention for production of proteins or polypeptides, through incorporation into cells, tissues, or organisms. In one embodiment, DNA containing all or part of the coding sequence for a Gene 216 polypeptide, or DNA which hybridizes to DNA having the sequence SEQ ID NO:1 or SEQ ID NO:6, is incorporated into a vector for expression of the encoded polypeptide in suitable host cells. The encoded polypeptide consisting of Gene 216, or its functional equivalent is capable of normal activity, such as proteolysis, adhesion, fusion, and intracellular activity.

The invention also concerns the use of the nucleotide sequence of the nucleic acids of this invention to identify DNA probes for Gene 216 genes, PCR primers to amplify Gene 216 genes, nucleotide polymorphisms in Gene 216

genes, and regulatory elements of the Gene 216 genes.

The nucleic acids of the present invention find use as primers and templates for the recombinant production of disorder-associated peptides or polypeptides, for chromosome and gene mapping, to provide antisense  
5 sequences, for tissue distribution studies, to locate and obtain full length genes, to identify and obtain homologous sequences (wild-type and mutants), and in diagnostic applications.

Probes may also be used for the detection of Gene 216-related sequences, and should preferably contain at least 50%, preferably at least  
10 80%, identity to Gene 216 polynucleotide, or a complementary sequence, or fragments thereof. The probes of this invention may be DNA or RNA, the probes may comprise all or a portion of the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:6, or a complementary sequence thereof, and may include  
15 promoter, enhancer elements, and introns of the naturally occurring Gene 216 polynucleotide.

The probes and primers based on the Gene 216 gene sequences disclosed herein are used to identify homologous Gene 216 gene sequences and proteins in other species. These Gene 216 gene sequences and proteins are used in the diagnostic/prognostic, therapeutic and drug-screening methods  
20 described herein for the species from which they have been isolated.

#### **Vectors and Host Cells**

The invention also provides vectors comprising the disorder-associated sequences, or derivatives or fragments thereof, and host cells for the production of purified proteins. A large number of vectors, including bacterial,  
25 yeast, and mammalian vectors, have been described for replication and/or expression in various host cells or cell-free systems, and may be used for gene therapy as well as for simple cloning or protein expression.

In one aspect, an expression vectors comprises a nucleic acid encoding a Gene 216 polypeptide or peptide, as described herein, operably linked to at  
30 least one regulatory sequence. Regulatory sequences are known in the art and are selected to direct expression of the desired protein in an appropriate

host cell. Accordingly, the term regulatory sequence includes promoters, enhancers and other expression control elements (see D.V. Goeddel (1990) *Methods Enzymol.* **185**:3-7). Enhancer and other expression control sequences are described in *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1983). It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transfected and/or the type of polypeptide desired to be expressed.

Several regulatory elements (e.g., promoters) have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. Such regulatory regions, methods of isolation, manner of manipulation, etc. are known in the art. Non-limiting examples of bacterial promoters include the  $\beta$ -lactamase (penicillinase) promoter; lactose promoter; tryptophan (trp) promoter; araBAD (arabinose) operon promoter; lambda-derived P<sub>1</sub> promoter and N gene ribosome binding site; and the hybrid tac promoter derived from sequences of the trp and lac UV5 promoters. Non-limiting examples of yeast promoters include the 3-phosphoglycerate kinase promoter, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter, galactokinase (GAL1) promoter, galactose epimerase promoter, and alcohol dehydrogenase (ADH1) promoter. Suitable promoters for mammalian cells include, without limitation, viral promoters, such as those from Simian Virus 40 (SV40), Rous sarcoma virus (RSV), adenovirus (ADV), and bovine papilloma virus (BPV). Preferred replication and inheritance systems include M13, ColE1, SV40, baculovirus, lambda, adenovirus, CEN ARS, 2 $\mu$ m ARS and the like. While expression vectors may replicate autonomously, they may also replicate by being inserted into the genome of the host cell, by methods well known in the art.

To obtain expression in eukaryotic cells, terminator sequences, polyadenylation sequences, and enhancer sequences that modulate gene expression may be required. Sequences that cause amplification of the gene may also be desirable. These sequences are well known in the art.

Furthermore, sequences that facilitate secretion of the recombinant product from cells, including, but not limited to, bacteria, yeast, and animal cells, such as secretory signal sequences and/or preprotein or proprotein sequences, may also be included. Such sequences are well described in the art.

5           Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector. The presence of this gene ensures growth of only those host cells that express the inserts. Typical selection genes encode proteins that 1) confer resistance to antibiotics or other toxic substances, e.g.  
10   ampicillin, neomycin, methotrexate, etc.; 2) complement auxotrophic deficiencies, or 3) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for Bacilli. Markers may be an inducible or non-inducible gene and will generally allow for positive selection.

Non-limiting examples of markers include the ampicillin resistance marker (i.e.,  
15   beta-lactamase), tetracycline resistance marker, neomycin/kanamycin resistance marker (i.e., neomycin phosphotransferase), dihydrofolate reductase, glutamine synthetase, and the like. The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts as understood by those of skill in the art.

20           Suitable expression vectors for use with the present invention include, but are not limited to, pUC, pBluescript (Stratagene), pET (Novagen, Inc., Madison, WI), and pREP (Invitrogen) plasmids. Vectors can contain one or more replication and inheritance systems for cloning or expression, one or more markers for selection in the host, e.g. antibiotic resistance, and one or  
25   more expression cassettes. The inserted coding sequences can be synthesized by standard methods, isolated from natural sources, or prepared as hybrids. Ligation of the coding sequences to transcriptional regulatory elements (e.g., promoters, enhancers, and/or insulators) and/or to other amino acid encoding sequences can be carried out using established methods.

30           Suitable cell-free expression systems for use with the present invention include, without limitation, rabbit reticulocyte lysate, wheat germ extract, canine

pancreatic microsomal membranes, *E. coli* S30 extract, and coupled transcription/translation systems (Promega Corp., Madison, WI). These systems allow the expression of recombinant polypeptides or peptides upon the addition of cloning vectors, DNA fragments, or RNA sequences containing  
5 protein-coding regions and appropriate promoter elements.

Non-limiting examples of suitable host cells include bacteria, archaea, insect, fungi (e.g., yeast), plant, and animal cells (e.g., mammalian, especially human). Of particular interest are *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, SF9 cells, C129 cells, 293 cells, *Neurospora*, and  
10 immortalized mammalian myeloid and lymphoid cell lines. Techniques for the propagation of mammalian cells in culture are well-known (see, Jakoby and Pastan (eds), 1979, *Cell Culture. Methods in Enzymology*, volume 58, Academic Press, Inc., Harcourt Brace Jovanovich, NY). Examples of commonly used mammalian host cell lines are VERO and HeLa cells, CHO  
15 cells, and WI38, BHK, and COS cell lines, although it will be appreciated by the skilled practitioner that other cell lines may be used, e.g., to provide higher expression desirable glycosylation patterns, or other features.

Host cells can be transformed, transfected, or infected as appropriate by any suitable method including electroporation, calcium chloride-, lithium  
20 chloride-, lithium acetate/polyethylene glycol-, calcium phosphate-, DEAE-dextran-, liposome-mediated DNA uptake, spheroplasting, injection, microinjection, microprojectile bombardment, phage infection, viral infection, or other established methods. Alternatively, vectors containing the nucleic acids of interest can be transcribed *in vitro*, and the resulting RNA introduced  
25 into the host cell by well-known methods, e.g., by injection (see, Kubo et al., 1988, *FEBS Letts.* **241**:119). The cells into which have been introduced nucleic acids described above are meant to also include the progeny of such cells.

The nucleic acids of the invention may be isolated directly from cells.  
30 Alternatively, the polymerase chain reaction (PCR) method can be used to produce the nucleic acids of the invention, using either RNA (e.g., mRNA) or

DNA (e.g., genomic DNA) as templates. Primers used for PCR can be synthesized using the sequence information provided herein and can further be designed to introduce appropriate new restriction sites, if desirable, to facilitate incorporation into a given vector for recombinant expression.

5           Using the information provided in SEQ ID NO:1 and SEQ ID NO:6, one skilled in the art will be able to clone and sequence all representative nucleic acids of interest, including nucleic acids encoding complete protein-coding sequences. It is to be understood that non-protein-coding sequences contained within SEQ ID NO:1 and SEQ ID NO:3 and the genomic sequences  
10 of SEQ ID NO:6 and SEQ ID NO:5 are also within the scope of the invention. Such sequences include, without limitation, sequences important for replication, recombination, transcription, and translation. Non-limiting examples include promoters and regulatory binding sites involved in regulation of gene expression, and 5'- and 3'- untranslated sequences (e.g., ribosome-  
15 binding sites) that form part of mRNA molecules.

The nucleic acids of this invention can be produced in large quantities by replication in a suitable host cell. Natural or synthetic nucleic acid fragments, comprising at least ten contiguous bases coding for a desired peptide or polypeptide can be incorporated into recombinant nucleic acid  
20 constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the nucleic acid constructs will be suitable for replication in a unicellular host, such as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cells, cell lines,  
25 tissues, or organisms. The purification of nucleic acids produced by the methods of the present invention is described, for example, in Sambrook et al., 1989; F.M. Ausubel et al., 1992, *Current Protocols in Molecular Biology*, J. Wiley and Sons, New York, NY.

The nucleic acids of the present invention can also be produced by  
30 chemical synthesis, e.g., by the phosphoramidite method described by Beaucage et al., 1981, *Tetra. Letts.* **22**:1859-1862, or the triester method

according to Matteucci et al., 1981, *J. Am. Chem. Soc.*, **103**:3185, and can performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and  
5 annealing the strands together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

These nucleic acids can encode full-length variant forms of proteins as well as the wild-type protein. The variant proteins (which could be especially  
10 useful for detection and treatment of disorders) will have the variant amino acid sequences encoded by the polymorphisms described in Table 10, when said polymorphisms are read so as to be in-frame with the full-length coding sequence of which it is a component.

Large quantities of the nucleic acids and proteins of the present  
15 invention may be prepared by expressing the Gene 216 nucleic acids or portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* or *Pseudomonas* may also be used. Mammalian or other  
20 eukaryotic host cells, such as those of yeast, filamentous fungi, plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. For example, insect cell systems (i.e., lepidopteran host cells and baculovirus expression vectors) are particularly suited for large-scale protein production.

25 Host cells carrying an expression vector (i.e., transformants or clones) are selected using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of  
30 a particular product based on temperature sensitivity may also serve as an appropriate marker.



Prokaryotic or eukaryotic cells comprising the nucleic acids of the present invention will be useful not only for the production of the nucleic acids and proteins of the present invention, but also, for example, in studying the characteristics of Gene 216 proteins. Cells and animals that carry the Gene 216 gene can be used as model systems to study and test for substances that have potential as therapeutic agents. The cells are typically cultured mesenchymal stem cells. These may be isolated from individuals with somatic or germline Gene 216 gene. Alternatively, the cell line can be engineered to carry the Gene 216 genes, as described above. After a test substance is applied to the cells, the transformed phenotype of the cell is determined. Any trait of transformed cells can be assessed, including respiratory diseases including asthma, atopy, and response to application of putative therapeutic agents.

#### **Antisense Nucleic Acids**

A further embodiment of the invention is antisense nucleic acids or oligonucleotides that are complementary, in whole or in part, to a target molecule comprising a sense strand of Gene 216. The Gene 216 target can be DNA, or its RNA counterpart (i.e., wherein thymine (T) is present in DNA and uracil (U) is present in RNA). When introduced into a cell, antisense nucleic acids or oligonucleotides can hybridize to all or a part of the sense strand of Gene 216, thereby inhibiting gene expression or replication.

In a particular embodiment of the invention, an antisense nucleic acid or oligonucleotide is wholly or partially complementary to, and can hybridize with, a target nucleic acid (either DNA or RNA) having the sequence of SEQ ID NO:1 or SEQ ID NO:6. For example, an antisense nucleic acid or oligonucleotide comprising 16 nucleotides can be sufficient to inhibit expression of the Gene 216 protein. Alternatively, an antisense nucleic acid or oligonucleotide can be complementary to 5' or 3' untranslated regions, or can overlap the translation initiation codon (5' untranslated and translated regions) of the Gene 216 gene, or its functional equivalent. In another embodiment, the antisense nucleic acid is wholly or partially complementary

to, and can hybridize with, a target nucleic acid that encodes a Gene 216 polypeptide.

In addition, oligonucleotides can be constructed which will bind to duplex nucleic acid (i.e., DNA:DNA or DNA:RNA), to form a stable triple helix-containing or triplex nucleic acid. Such triplex oligonucleotides can inhibit transcription and/or expression of a gene encoding Gene 216, or its functional equivalent (M.D. Frank-Kamenetskii and S.M. Mirkin, 1995, *Ann. Rev. Biochem.* **64**:65-95). Triplex oligonucleotides are constructed using the base-pairing rules of triple helix formation and the nucleotide sequence of the gene or mRNA for Gene 216.

The present invention encompasses methods of using oligonucleotides in antisense inhibition of the function of Gene 216. In the context of this invention, the term "oligonucleotide" refers to naturally-occurring species or synthetic species formed from naturally-occurring subunits or their close homologs. The term may also refer to moieties that function similarly to oligonucleotides, but have non-naturally-occurring portions. Thus, oligonucleotides may have altered sugar moieties or inter-sugar linkages. Exemplary among these are phosphorothioate and other sulfur containing species which are known in the art.

In preferred embodiments, at least one of the phosphodiester bonds of the oligonucleotide has been substituted with a structure that functions to enhance the ability of the compositions to penetrate into the region of cells where the RNA whose activity is to be modulated is located. It is preferred that such substitutions comprise phosphorothioate bonds, methyl phosphonate bonds, or short chain alkyl or cycloalkyl structures. In accordance with other preferred embodiments, the phosphodiester bonds are substituted with structures which are, at once, substantially non-ionic and non-chiral, or with structures which are chiral and enantiomerically specific. Persons of ordinary skill in the art will be able to select other linkages for use in the practice of the invention.

Oligonucleotides may also include species that include at least some

modified base forms. Thus, purines and pyrimidines other than those normally found in nature may be so employed. Similarly, modifications on the furanosyl portions of the nucleotide subunits may also be effected, as long as the essential tenets of this invention are adhered to. Examples of such modifications are 2'-O-alkyl- and 2'-halogen-substituted nucleotides. Some non-limiting examples of modifications at the 2' position of sugar moieties which are useful in the present invention include OH, SH, SCH<sub>3</sub>, F, OCH<sub>3</sub>, OCN, O(CH<sub>2</sub>)<sub>n</sub> NH<sub>2</sub> and O(CH<sub>2</sub>)<sub>n</sub> CH<sub>3</sub>, where n is from 1 to about 10. Such oligonucleotides are functionally interchangeable with natural oligonucleotides or synthesized oligonucleotides, which have one or more differences from the natural structure. All such analogs are comprehended by this invention so long as they function effectively to hybridize with Gene 216 DNA or RNA to inhibit the function thereof.

The oligonucleotides in accordance with this invention preferably comprise from about 3 to about 50 subunits. It is more preferred that such oligonucleotides and analogs comprise from about 8 to about 25 subunits and still more preferred to have from about 12 to about 20 subunits. As defined herein, a "subunit" is a base and sugar combination suitably bound to adjacent subunits through phosphodiester or other bonds.

Antisense nucleic acids or oligonucleotides can be produced by standard techniques (see, e.g., Shewmaker et al., U.S. Patent No. 5,107,065. The oligonucleotides used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is available from several vendors, including PE Applied Biosystems (Foster City, CA). Any other means for such synthesis may also be employed, however, the actual synthesis of the oligonucleotides is well within the abilities of the practitioner. It is also known to prepare other oligonucleotide such as phosphorothioates and alkylated derivatives.

The oligonucleotides of this invention are designed to be hybridizable with Gene 216 RNA (e.g., mRNA) or DNA. For example, an oligonucleotide

(e.g., DNA oligonucleotide) that hybridizes to Gene 216 mRNA can be used to target the mRNA for RnaseH digestion. Alternatively, an oligonucleotide that hybridizes to the translation initiation site of Gene 216 mRNA can be used to prevent translation of the mRNA. In another approach, oligonucleotides that bind to the double-stranded DNA of Gene 216 can be administered. Such oligonucleotides can form a triplex construct and inhibit the transcription of the DNA encoding Gene 216 polypeptides. Triple helix pairing prevents the double helix from opening sufficiently to allow the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described (see, e.g., J.E. Gee et al., 1994, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY).

As non-limiting examples, antisense oligonucleotides may be targeted to hybridize to the following regions: mRNA cap region; translation initiation site; translational termination site; transcription initiation site; transcription termination site; polyadenylation signal; 3' untranslated region; 5' untranslated region; 5' coding region; mid coding region; and 3' coding region. Preferably, the complementary oligonucleotide is designed to hybridize to the most unique 5' sequence Gene 216, including any of about 15-35 nucleotides spanning the 5' coding sequence. Appropriate oligonucleotides can be designed using OLIGO software (Molecular Biology Insights, Inc., Cascade, CO; <http://www.oligo.net>).

In accordance with the present invention, the antisense oligonucleotide can be synthesized, formulated as a pharmaceutical composition, and administered to a subject. The synthesis and utilization of antisense and triplex oligonucleotides have been previously described (e.g., H. Simon et al., 1999, *Antisense Nucleic Acid Drug Dev.* **9**:527-31; F.X. Barre et al., 2000, *Proc. Natl. Acad. Sci. USA* **97**:3084-3088; R. Elez et al., 2000, *Biochem. Biophys. Res. Commun.* **269**:352-6; E.R. Sauter et al., 2000, *Clin. Cancer Res.* **6**:654-60). Alternatively, expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population.

Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express nucleic acid sequence that is complementary to the nucleic acid sequence encoding a Gene 216 polypeptide. These techniques are described both in Sambrook et al., 1989 and in Ausubel et al., 1992. For example, Gene 216 expression can be inhibited by transforming a cell or tissue with an expression vector that expresses high levels of untranslatable sense or antisense Gene 216 sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and even longer if appropriate replication elements included in the vector system.

Various assays may be used to test the ability of Gene 216-specific antisense oligonucleotides to inhibit Gene 216 expression. For example, Gene 216 mRNA levels can be assessed northern blot analysis (Sambrook et al., 1989; Ausubel et al., 1992; J.C. Alwine et al. 1977, *Proc. Natl. Acad. Sci. USA* **74**:5350-5354; I.M. Bird, 1998, *Methods Mol. Biol.* **105**:325-36), quantitative or semi-quantitative RT-PCR analysis (see, e.g., W.M. Freeman et al., 1999, *Biotechniques* **26**:112-122; Ren et al., 1998, *Mol. Brain Res.* **59**:256-63; J.M. Cale et al., 1998, *Methods Mol. Biol.* **105**:351-71), or *in situ* hybridization (reviewed by A.K. Raap, 1998, *Mutat. Res.* **400**:287-298). Alternatively, antisense oligonucleotides may be assessed by measuring levels of Gene 216 polypeptide, e.g., by western blot analysis, indirect immunofluorescence, immunoprecipitation techniques (see, e.g., J.M. Walker, 1998, *Protein Protocols on CD-ROM*, Humana Press, Totowa, NJ).

### **Polypeptides**

The invention also relates to polypeptides and peptides encoded by the novel nucleic acids described herein. The polypeptides and peptides of this invention can be isolated and/or recombinant. In a preferred embodiment, the Gene 216 polypeptide, or analog or portion thereof, has at least one function characteristic of a Gene 216 protein, for example, proteolysis, adhesion,

fusion, antigenic, and intracellular activity. Protein analogs include, for example, naturally-occurring or genetically engineered Gene 216 variants (e.g. mutants) and portions thereof. Variants may differ from wild-type Gene 216 protein by the addition, deletion, or substitution of one or more amino acid residues. In specific embodiments, polypeptide variants are encoded by Gene 216 nucleic acids containing one or more of the SNPs disclosed herein. Variants also include polypeptides in which one or more residues are modified (i.e., by phosphorylation, sulfation, acylation, etc.), and mutants comprising one or more modified residues.

Variant polypeptides can have conservative changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More infrequently, a variant polypeptide can have non-conservative changes, e.g., substitution of a glycine with a tryptophan. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, for example, DNASTAR software (DNASTAR, Inc., Madison, WI)

As non-limiting examples, conservative substitutions in the Gene 216 amino acid sequence can be made in accordance with the following table:

Original Residue	Conservative Substitution(s)
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser

Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

Substantial changes in function or immunogenicity can be made by selecting substitutions that are less conservative than those shown in the table, above. For example, non-conservative substitutions can be made which more significantly affect the structure of the polypeptide in the area of the alteration, for example, the alpha-helical, or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which generally are expected to produce the greatest changes in the polypeptide's properties are those where 1) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl, or alanyl; 2) a cysteine or proline is substituted for (or by) any other residue; 3) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or 4) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) a residue that does not have a side chain, e.g., glycine.

In one embodiment, polypeptides of the present invention share at least 50% amino acid sequence identity with a Gene 216 polypeptide, such as SEQ ID NO:4, or fragments thereof. Preferably, the polypeptides share at least 65% amino acid sequence identity; more preferably, the polypeptides share at least 75% amino acid sequence identity; even more preferably, the polypeptides share at least 80% amino acid sequence identity with a Gene 216 polypeptide; still more preferably the polypeptides share at least 90% amino acid sequence identity with a Gene 216 polypeptide.

Percent sequence identity can be calculated using computer programs or direct sequence comparison. Preferred computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package, FASTA, BLASTP, and TBLASTN (see, e.g., D.W. Mount, 2001, *Bioinformatics: Sequence and Genome Analysis*, Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, NY). The BLASTP and TBLASTN programs are publicly available from NCBI and other sources. The well-known Smith Waterman algorithm may also be used to determine identity.

Exemplary parameters for amino acid sequence comparison include the following: 1) algorithm from Needleman and Wunsch, 1970, *J Mol. Biol.* 48:443-453; 2) BLOSSUM62 comparison matrix from Hentikoff and Hentikoff, 1992, *Proc. Natl. Acad. Sci. USA* 89:10915-10919; 3) gap penalty = 12; and 4) gap length penalty = 4. A program useful with these parameters is publicly available as the "gap" program (Genetics Computer Group, Madison, WI). The aforementioned parameters are the default parameters for polypeptide comparisons (with no penalty for end gaps).

Alternatively, polypeptide sequence identity can be calculated using the following equation: % identity = (the number of identical residues) / (alignment length in amino acid residues) \* 100. For this calculation, alignment length includes internal gaps but does not include terminal gaps.

In accordance with the present invention, polypeptide sequences may be identical to the sequence of SEQ ID NO:4, or may include up to a certain integer number of amino acid alterations. Polypeptide alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. Alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In specific embodiments, polypeptide variants may be encoded by Gene 216 nucleic acids comprising SNPs and/or alternate splice variants.

The invention also relates to isolated, synthesized and/or recombinant portions or fragments of a Gene 216 protein or polypeptide as described herein. Polypeptide fragments (i.e., peptides) can be made which have full or partial function on their own, or which when mixed together (though fully, partially, or nonfunctional alone), spontaneously assemble with one or more



other polypeptides to reconstitute a functional protein having at least one functional characteristic of a Gene 216 protein of this invention. In addition, Gene 216 polypeptide fragments may comprise, for example, one or more domains of the Gene 216 polypeptide (e.g., the pre-, pro-, catalytic, cysteine-rich, disintegrin, EGF, transmembrane, and cytoplasmic domains) disclosed  
5 herein.

Polypeptides according to the invention can comprise at least 5 amino acid residues; preferably the polypeptides comprise at least 12 residues; more preferably the polypeptides comprise at least 20 residues; and yet more  
10 preferably the polypeptides comprise at least 30 residues. Nucleic acids comprising protein-coding sequences can be used to direct the expression of asthma-associated polypeptides in intact cells or in cell-free translation systems. The coding sequence can be tailored, if desired, for more efficient expression in a given host organism, and can be used to synthesize  
15 oligonucleotides encoding the desired amino acid sequences. The resulting oligonucleotides can be inserted into an appropriate vector and expressed in a compatible host organism or translation system.

The polypeptides of the present invention, including function-conservative variants, may be isolated from wild-type or mutant cells (e.g.,  
20 human cells or cell lines), from heterologous organisms or cells (e.g., bacteria, yeast, insect, plant, and mammalian cells), or from cell-free translation systems (e.g., wheat germ, microsomal membrane, or bacterial extracts) in which a protein-coding sequence has been introduced and expressed. Furthermore, the polypeptides may be part of recombinant fusion proteins. The polypeptides  
25 can also, advantageously, be made by synthetic chemistry. Polypeptides may be chemically synthesized by commercially available automated procedures, including, without limitation, exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis.

Methods for polypeptide purification are well-known in the art, including,  
30 without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition

chromatography, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide in a recombinant system in which the protein contains an additional sequence (e.g., epitope or protein) tag that facilitates purification. Non-limiting examples of epitope tags include c-myc, 5 haemagglutinin (HA), polyhistidine (6X-HIS) (SEQ ID NO:32), GLU-GLU, and DYKDDDDK (SEQ ID NO:33) (FLAG®) epitope tags. Non-limiting examples of protein tags include glutathione-S-transferase (GST), green fluorescent protein (GFP), and maltose binding protein (MBP).

In one approach, the coding sequence of a polypeptide or peptide can be cloned into a vector that creates a fusion with a sequence tag of interest. 10 Suitable vectors include, without limitation, pRSET (Invitrogen Corp., San Diego, CA), pGEX (Amersham-Pharmacia Biotech, Inc., Piscataway, NJ), pEGFP (CLONTECH Laboratories, Inc., Palo Alto, CA), and pMAL™ (New England BioLabs (NEB), Inc., Beverly, MA) plasmids. Following expression, 15 the epitope, or protein tagged polypeptide or peptide can be purified from a crude lysate of the translation system or host cell by chromatography on an appropriate solid-phase matrix. In some cases, it may be preferable to remove the epitope or protein tag (i.e., via protease cleavage) following purification. As an alternative approach, antibodies produced against a disorder-associated 20 protein or against peptides derived therefrom can be used as purification reagents. Other purification methods are possible.

The present invention also encompasses polypeptide derivatives of Gene 216. The isolated polypeptides may be modified by, for example, phosphorylation, sulfation, acylation, or other protein modifications. They may 25 also be modified with a label capable of providing a detectable signal, either directly or indirectly, including, but not limited to, radioisotopes and fluorescent compounds.

Both the *naturally occurring and recombinant* forms of the polypeptides of the invention can advantageously be used to screen compounds for binding 30 activity. Many methods of screening for binding activity are known by those skilled in the art and may be used to practice the invention. Several methods

of automated assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period of time. Such high-throughput screening methods are particularly preferred. The use of high-throughput screening assays to test for inhibitors is greatly facilitated by the availability of large amounts of purified polypeptides, as provided by the invention. The polypeptides of the invention also find use as therapeutic agents as well as antigenic components to prepare antibodies.

The polypeptides of this invention find use as immunogenic components useful as antigens for preparing antibodies by standard methods. It is well known in the art that immunogenic epitopes generally contain at least about five amino acid residues (Ohno et al., 1985, *Proc. Natl. Acad. Sci. USA* 82:2945). Therefore, the immunogenic components of this invention will typically comprise at least 5 amino acid residues of the sequence of the complete polypeptide chains. Preferably, they will contain at least 7, and most preferably at least about 10 amino acid residues or more to ensure that they will be immunogenic. Whether a given component is immunogenic can readily be determined by routine experimentation. Such immunogenic components can be produced by proteolytic cleavage of larger polypeptides or by chemical synthesis or recombinant technology and are thus not limited by proteolytic cleavage sites. The present invention thus encompasses antibodies that specifically recognize asthma-associated immunogenic components.

### **Structural Studies**

A purified Gene 216 polypeptide can be analyzed by well-established methods (e.g., X-ray crystallography, NMR, CD, etc.) to determine the three-dimensional structure of the molecule. The three-dimensional structure, in turn, can be used to model intermolecular interactions. Exemplary methods for crystallization and X-ray crystallography are found in P.G. Jones, 1981, *Chemistry in Britain*, 17:222-225; C. Jones et al. (eds), *Crystallographic Methods and Protocols*, Humana Press, Totowa, NJ; A. McPherson, 1982, *Preparation and Analysis of Protein Crystals*, John Wiley & Sons, New York, NY; T.L. Blundell and L.N. Johnson, 1976, *Protein Crystallography*, Academic

Press, Inc., New York, NY; A. Holden and P. Singer, 1960, *Crystals and Crystal Growing*, Anchor Books-Doubleday, New York, NY; R.A. Laudise, 1970, *The Growth of Single Crystals*, Solid State Physical Electronics Series, N. Holonyak, Jr., (ed), Prentice-Hall, Inc.; G.H. Stout and L.H. Jensen, 1989, X-ray Structure Determination: A Practical Guide, 2nd edition, John Wiley & Sons, New York, NY; *Fundamentals of Analytical Chemistry*, 3rd. edition, Saunders Golden Sunburst Series, Holt, Rinehart and Winston, Philadelphia, PA, 1976; P.D. Boyle of the Department of Chemistry of North Carolina State University at <http://laue.chem.ncsu.edu/web/GrowXtal.html>; M.B. Berry, 1995, *Protein Crystalization: Theory and Practice, Structure and Dynamics of E. coli Adenylate Kinase*, Doctoral Thesis, Rice University, Houston TX; [www.bioc.rice.edu/~berry/papers/crystalization/crystalization.html](http://www.bioc.rice.edu/~berry/papers/crystalization/crystalization.html).

For X-ray diffraction studies, single crystals can be grown to suitable size. Preferably, a crystal has a size of 0.2 to 0.4 mm in at least two of the three dimensions. Crystals can be formed in a solution comprising a Gene 216 polypeptide (e.g., 1.5-200 mg/ml) and reagents that reduce the solubility to conditions close to spontaneous precipitation. Factors that affect the formation of polypeptide crystals include: 1) purity; 2) substrates or co-factors; 3) pH; 4) temperature; 5) polypeptide concentration; and 6) characteristics of the precipitant. Preferably, the Gene 216 polypeptides are pure, i.e., free from contaminating components (at least 95% pure), and free from denatured Gene 216 polypeptides. In particular, polypeptides can be purified by FPLC and HPLC techniques to assure homogeneity (see, Lin et al., 1992, *J. Crystal Growth*. **122**:242-245). Optionally, Gene 216 polypeptide substrates or co-factors can be added to stabilize the quaternary structure of the protein and promote lattice packing.

Suitable precipitants for crystallization include, but are not limited to, salts (e.g., ammonium sulphate, potassium phosphate); polymers (e.g., polyethylene glycol (PEG) 6000); alcohols (e.g., ethanol); polyalcohols (e.g., 1-methyl-2,4 pentane diol (MPD)); organic solvents; sulfonic dyes; and deionized water. The ability of a salt to precipitate polypeptides can be

generally described by the Hofmeister series:  $\text{PO}_4^{3-} > \text{HPO}_4^{2-} = \text{SO}_4^{2-} > \text{citrate} > \text{CH}_3\text{CO}_2^- > \text{Cl}^- > \text{Br}^- > \text{NO}_3^- > \text{ClO}_4^- > \text{SCN}^-$ ; and  $\text{NH}_4^+ > \text{K}^+ > \text{Na}^+ > \text{Li}^+$ . Non-limiting examples of salt precipitants are shown below (see Berry, 1995).

Precipitant	Maximum concentration
$(\text{NH}_4^+/\text{Na}^+/\text{Li}^+)_2$ or $\text{Mg}_2+\text{SO}_4^{2-}$	4.0 / 1.5 / 2.1 / 2.5 M
$\text{NH}_4^+/\text{Na}^+/\text{K}^+ \text{PO}_4^{3-}$	3.0 / 4.0 / 4.0 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+$ citrate	~1.8 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+$ acetate	~3.0 M
$\text{NH}_4^+/\text{K}^+/\text{Na}^+/\text{Li}^+ \text{Cl}^-$	5.2 / 9.8 / 4.2 / 5.4 M
$\text{NH}_4^+ \text{NO}_3^-$	~8.0 M

- 5 High molecular weight polymers useful as precipitating agents include polyethylene glycol (PEG), dextran, polyvinyl alcohol, and polyvinyl pyrrolidone (A. Polson et al., 1964, *Biochem. Biophys. Acta.* **82**:463-475). In general, polyethylene glycol (PEG) is the most effective for forming crystals. PEG compounds with molecular weights less than 1000 can be used at  
10 concentrations above 40% v/v. PEGs with molecular weights above 1000 can be used at concentration 5-50% w/v. Typically, PEG solutions are mixed with ~0.1 % sodium azide to prevent bacterial growth.

- Typically, crystallization requires the addition of buffers and a specific salt content to maintain the proper pH and ionic strength for a protein's stability.  
15 Suitable additives include, but are not limited to sodium chloride (e.g., 50-500 mM as additive to PEG and MPD; 0.15-2 M as additive to PEG); potassium chloride (e.g., 0.05-2 M); lithium chloride (e.g., 0.05-2 M); sodium fluoride (e.g., 20-300 mM); ammonium sulfate (e.g., 20-300 mM); lithium sulfate (e.g., 0.05-2 M); sodium or ammonium thiocyanate (e.g., 50-500 mM); MPD (e.g., 0.5-50%);  
20 1,6 hexane diol (e.g., 0.5-10%); 1,2,3 heptane triol (e.g., 0.5-15%); and benzamidine (e.g., 0.5-15%).

- Detergents may be used to maintain protein solubility and prevent aggregation. Suitable detergents include, but are not limited to non-ionic detergents such as sugar derivatives, oligoethyleneglycol derivatives,  
25 dimethylamine-N-oxides, cholate derivatives, N-octyl hydroxyalkylsulphoxides, sulphobetains, and lipid-like detergents. Sugar-derived detergents include alkyl glucopyranosides (e.g., C8-GP, C9-GP), alkyl thio-glucopyranosides (e.g., C8-

tGP), alkyl maltopyranosides (e.g., C10-M, C12-M; CYMAL-3, CYMAL-5, CYMAL-6), alkyl thio-maltopyranosides, alkyl galactopyranosides, alkyl sucroses (e.g., N-octanoylsucrose), and glucamides (e.g., HECAMEG, C-HEGA-10; MEGA-8). Oligoethyleneglycol-derived detergents include alkyl  
5 polyoxyethylenes (e.g., C8-E5, C8-En; C12-E8; C12-E9) and phenyl polyoxyethylenes (e.g., Triton X-100). Dimethylamine-N-oxide detergents include, e.g., C10-DAO; DDAO; LDAO. Cholate-derived detergents include, e.g., Deoxy-Big CHAP, digitonin. Lipid-like detergents include phosphocholine compounds. Suitable detergents further include zwitter-ionic detergents (e.g.,  
10 ZWITTERGENT 3-10; ZWITTERGENT 3-12); and ionic detergents (e.g., SDS).

Crystallization of macromolecules has been performed at temperatures ranging from 60°C to less than 0°C. However, most molecules can be crystallized at 4°C or 22°C. Lower temperatures promote stabilization of polypeptides and inhibit bacterial growth. In general, polypeptides are more  
15 soluble in salt solutions at lower temperatures (e.g., 4°C), but less soluble in PEG and MPD solutions at lower temperatures. To allow crystallization at 4°C or 22°C, the precipitant or protein concentration can be increased or decreased as required. Heating, melting, and cooling of crystals or aggregates can be used to enlarge crystals. In addition, crystallization at both 4°C and 22°C can  
20 be assessed (A. McPherson, 1992, *J. Cryst. Growth*. **122**:161-167; C.W. Carter, Jr. and C.W. Carter, 1979, *J. Biol. Chem.* **254**:12219-12223; T. Bergfors, 1993, *Crystallization Lab Manual*).

A crystallization protocol can be adapted to a particular polypeptide or peptide. In particular, the physical and chemical properties of the polypeptide  
25 can be considered (e.g., aggregation, stability, adherence to membranes or tubing, internal disulfide linkages, surface cysteines, chelating ions, etc.). For initial experiments, the standard set of crystallization reagents can be used (Hampton Research, Laguna Niguel, CA). In addition, the CRYSTOOL program can provide guidance in determining optimal crystallization conditions  
30 (Brent Segelke, 1995, Efficiency analysis of sampling protocols used in protein crystallization screening and crystal structure from two novel crystal forms of

PLA2, Ph.D. Thesis, University of California, San Diego; <http://www.ccp14.ac.uk/ccp/web-mirrors/linrupp/crystool/crystool.htm>). Exemplary crystallization conditions are shown below (see Berry, 1995).

Major Precipitant	Additive	Concentration of Major Precipitant	Concentration of Additive
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	PEG 400-2000, MPD, ethanol, or methanol	2.0-4.0 M	6%-0.5%
Na citrate	PEG 400-2000, MPD, ethanol, or methanol	1.4-1.8 M	6%-0.5%
PEG 1000-20000	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , NaCl, or Na formate	40-50%	0.2-0.6 M

5

Robots can be used for automatic screening and optimization of crystallization conditions. For example, the IMPAX and Oryx systems can be used (Douglas Instruments, Ltd., East Garston, United Kingdom). The CRYSTOOL program (Segelke, *supra*) can be integrated with the robotics programming. In addition, the Xact program can be used to construct, maintain, and record the results of various crystallization experiments (see, e.g., D.E. Brodersen et al., 1999, *J. Appl. Cryst.* **32**: 1012-1016; G.R. Andersen and J. Nyborg, 1996, *J. Appl. Cryst.* **29**:236-240). The Xact program supports multiple users and organizes the results of crystallization experiments into hierarchies. Advantageously, Xact is compatible with both CRYSTOOL and Microsoft® Excel programs.

Four methods are commonly employed to crystallize macromolecules: vapor diffusion, free interface diffusion, batch, and dialysis. The vapor diffusion technique is typically performed by formulating a 1:1 mixture of a solution comprising the polypeptide of interest and a solution containing the precipitant at the final concentration that is to be achieved after vapor equilibration. The drop containing the 1:1 mixture of protein and precipitant is then suspended and sealed over the well solution, which contains the precipitant at the target concentration, as either a hanging or sitting drop. Vapor diffusion can be used to screen a large number of crystallization conditions or when small amounts of polypeptide are available. For screening, drop sizes of 1 to 2  $\mu$ l can be used. Once preliminary crystallization conditions

have been determined, drop sizes such as 10  $\mu$ l can be used. Notably, results from hanging drops may be improved with agarose gels (see K. Provost and M.-C. Robert, 1991, *J. Cryst. Growth*. **110**:258-264). Free interface diffusion is performed by layering of a low density solution onto one of higher density, usually in the form of concentrated protein onto concentrated salt. Since the solute to be crystallized must be concentrated, this method typically requires relatively large amounts of protein. However, the method can be adapted to work with small amounts of protein. In a representative experiment, 2 to 5  $\mu$ l of sample is pipetted into one end of a 20  $\mu$ l microcapillary pipet. Next, 2 to 5  $\mu$ l of precipitant is pipetted into the capillary without introducing an air bubble, and the ends of the pipet are sealed. With sufficient amounts of protein, this method can be used to obtain relatively large crystals (see, e.g., S.M. Althoff et al., 1988, *J. Mol. Biol.* **199**:665-666).

The batch technique is performed by mixing concentrated polypeptide with concentrated precipitant to produce a final concentration that is supersaturated for the solute macromolecule. Notably, this method can employ relatively large amounts of solution (e.g., milliliter quantities), and can produce large crystals. For that reason, the batch technique is not recommended for screening initial crystallization conditions.

The dialysis technique is performed by diffusing precipitant molecules through a semipermeable membrane to slowly increase the concentration of the solute inside the membrane. Dialysis tubing can be used to dialyze milliliter quantities of sample, whereas dialysis buttons can be used to dialyze microliter quantities (e.g., 7-200  $\mu$ l). Dialysis buttons may be constructed out of glass, perspex, or Teflon™ (see, e.g., Cambridge Repetition Engineers Ltd., Greens Road, Cambridge CB4 3EQ, UK; Hampton Research). Using this method, the precipitating solution can be varied by moving the entire dialysis button or sack into a different solution. In this way, polypeptides can be "reused" until the correct conditions for crystallization are found (see, e.g., C.W. Carter, Jr. et al., 1988, *J. Cryst. Growth*. **90**:60-73). However, this method is not recommended for precipitants comprising concentrated PEG solutions.



Various strategies have been designed to screen crystallization conditions, including 1) pl screening; 2) grid screening; 3) factorials; 4) solubility assays; 5) perturbation; and 6) sparse matrices. In accordance with the pl screening method, the pl of a polypeptide is presumed to be its crystallization point. Screening at the pl can be performed by dialysis against low concentrations of buffer (less than 20 mM) at the appropriate pH, or by use of conventional precipitants.

The grid screening method can be performed on two-dimensional matrices. Typically, the precipitant concentration is plotted against pH. The optimal conditions can be determined for each axis, and then combined. At that point, additional factors can be tested (e.g., temperature, additives). This method works best with fast-forming crystals, and can be readily automated (see M.J. Cox and P.C. Weber, 1988, *J. Cryst. Growth*. **90**:318-324). Grid screens are commercially available for popular precipitants such as ammonium sulphate, PEG 6000, MPD, PEG/LiCl, and NaCl (see, e.g., Hamilton Research).

The incomplete factorial method can be performed by 1) selecting a set of ~20 conditions; 2) randomly assigning combinations of these conditions; 3) grading the success of the results of each experiment using an objective scale; and 4) statistically evaluating the effects of each of the conditions on crystal formation (see, e.g., C.W. Carter, Jr. et al., 1988, *J. Cryst. Growth*. **90**:60-73).

In particular, conditions such as pH, temperature, precipitating agent, and cations can be tested. Dialysis buttons are preferably used with this method. Typically, optimal conditions/combinations can be determined within 35 tests. Similar approaches, such as "footprinting" conditions, may also be employed (see, e.g., E.A. Stura et al., 1991, *J. Cryst. Growth*. **110**:1-2).

The perturbation approach can be performed by altering crystallization conditions by introducing a series of additives designed to test the effects of altering the structure of bulk solvent and the solvent dielectric on crystal formation (see, e.g., Whitaker et al., 1995, *Biochem*. **34**:8221-8226). Additives for increasing the solvent dielectric include, but are not limited to, NaCl, KCl,

or LiCl (e.g., 200 mM); Na formate (e.g., 200 mM); Na<sub>2</sub>HPO<sub>4</sub> or K<sub>2</sub>HPO<sub>4</sub> (e.g., 200 mM); urea, trichloroacetate, guanidium HCl, or KSCN (e.g., 20-50 mM).

A non-limiting list of additives for decreasing the solvent dielectric include methanol, ethanol, isopropanol, or tert-butanol (e.g., 1-5%); MPD (e.g., 1%);

- 5 PEG 400, PEG 600, or PEG 1000 (e.g., 1-4%); PEG MME (monomethylether) 550, PEG MME 750, PEG MME 2000 (e.g., 1-4%).

As an alternative to the above-screening methods, the sparse matrix approach can be used (see, e.g., J. Jancarik and S.-H.J. Kim, 1991, *Appl. Cryst.* **24**:409-411; A. McPherson, 1992, *J. Cryst. Growth.* **122**:161-167; B. Cudney et al., 1994, *Acta. Cryst.* **D50**:414-423). Sparse matrix screens are commercially available (see, e.g., Hampton Research; Molecular Dimensions, Inc., Apopka, FL; Emerald Biostructures, Inc., Lemont, IL). Notably, data from Hampton Research sparse matrix screens can be stored and analyzed using ASPRUN software (Douglas Instruments).

- 15 Exemplary conditions for an initial screen are shown below (see Berry, 1995).

**TABLE 1**

Tray 1:

PEG 8000 (wells 1-6)						Ammonium sulfate (wells 7-12)					
1	2	3	4	5	6	7	8	9	10	11	12
20% pH 5.0	20% pH 7.0	20% pH 8.6	35% pH 5.0	35% pH 7.0	35% pH 8.6	2.0 M pH 5.0	2.0 M pH 7.0	2.0 M pH 8.8	2.5 M pH 5.0	2.5 M pH 7.0	2.5 M pH 8.8
MPD (wells 13-16)				Na Citrate (wells 17-20)				Na/K Phosphate (wells 21-24)			
13	14	15	16	17	18	19	20	21	22	23	24
30% pH 5.8	30% pH 7.6	50% pH 5.8	50% pH 7.6	1.3 M pH 5.8	1.3 M pH 7.5	1.5 M pH 5.8	1.5 M pH 7.5	2.0 M pH 6.0	2.0 M pH 7.4	2.5 M pH 6.0	2.5 M pH 7.4

- 20 Tray 2:

PEG 2000 MME/0.2 M Ammon. sulfate (wells 25-30)					
25	26	27	28	29	30
25% pH 5.5	25% pH 7.0	25% pH 8.5	40% pH 6.5	40% pH 7.0	40% pH 8.5
Random for wells 31 to 48					

- The initial screen can be used with hanging or sitting drops. To conserve the sample, tray 2 can be set up several weeks following tray 1. Wells 31-48 of tray 2 can comprise a random set of solutions. Alternatively, solutions can be formulated using sparse methods. Preferably, test solutions cover a broad range of precipitants, additives, and pH (especially pH 5.0-9.0).

Seeding can be used to trigger nucleation and crystal growth (Stura and

Wilson, 1990, *J. Cryst. Growth*. **110**:270-282; C. Thaller et al., 1981, *J. Mol. Biol.* **147**:465-469; A. McPherson and P. Schlichta, 1988, *J. Cryst. Growth*. **90**:47-50). In general, seeding can be performed by transferring crystal seeds into a polypeptide solution to allow polypeptide molecules to deposit on the surface of the seeds and produce crystals. Two seeding methods can be used: microseeding and macroseeding. For microseeding, a crystal can be ground into tiny pieces and transferred into the protein solution. Alternatively, seeds can be transferred by adding 1-2  $\mu$ l of the seed solution directly to the equilibrated protein solution. In another approach, seeds can be transferred by dipping a hair in the seed solution and then streaking the hair across the surface of the drop (streak seeding; see Stura and Wilson, *supra*). For macroseeding, an intact crystal can be transferred into the protein solution (see, e.g., C. Thaller et al., 1981, *J. Mol. Biol.* **147**:465-469). Preferably, the surface of the crystal seed is washed to regenerate the growing surface prior to being transferred. Optimally, the protein solution for crystallization is close to saturation and the crystal seed is not completely dissolved upon transfer.

### **Antibodies**

An isolated Gene 216 polypeptide or a portion or fragment thereof, can be used as an immunogen to generate anti-Gene 216 antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length Gene 216 polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments of Gene 216 for use as immunogens. The antigenic peptide of Gene 216 comprises at least 5 amino acid residues of the amino acid sequence shown in SEQ ID NO:4, and encompasses an epitope of Gene 216 such that an antibody raised against the peptide forms a specific immune complex with Gene 216 amino acid sequence.

Accordingly, another aspect of the invention pertains to anti-Gene 216 antibodies. The invention provides polyclonal and monoclonal antibodies that bind Gene 216 polypeptides or peptides. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site

capable of immunoreacting with a particular epitope of a Gene 216 polypeptide or peptide. A monoclonal antibody composition thus typically displays a single binding affinity for a particular Gene 216 polypeptide or peptide with which it immunoreacts.

- 5           A Gene 216 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse, or other non-human mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed Gene 216 polypeptide or a chemically synthesized Gene 216 polypeptide, or fragments thereof. The  
10       preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic Gene 216 preparation induces a polyclonal anti-Gene 216 antibody response.

- A number of adjuvants are known and used by those skilled in the art.
- 15       Non-limiting examples of suitable adjuvants include incomplete Freund's adjuvant, mineral gels such as alum, aluminum phosphate, aluminum hydroxide, aluminum silica, and surface-active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Further examples of adjuvants include N-  
20       acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3  
      hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria,  
25       monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. A particularly useful adjuvant comprises 5% (wt/vol) squalene, 2.5% Pluronic L121 polymer and 0.2% polysorbate in phosphate buffered saline (Kwak et al., 1992, *New Eng.*  
      *J. Med.* **327**:1209-1215). Preferred adjuvants include complete BCG, Detox,  
30       (RIBI, Immunochem Research Inc.), ISCOMS, and aluminum hydroxide adjuvant (Superphos, Biosector). The effectiveness of an adjuvant may be

determined by measuring the amount of antibodies directed against the immunogenic peptide.

Polyclonal anti-Gene 216 antibodies can be prepared as described above by immunizing a suitable subject with a Gene 216 immunogen. The  
5 anti-Gene 216 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized Gene 216. If desired, the antibody molecules directed against Gene 216 can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A  
10 chromatography to obtain the IgG fraction.

At an appropriate time after immunization, e.g., when the anti-Gene 216 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique (see Kohler and Milstein, 1975, *Nature*  
15 **256**:495-497; Brown et al., 1981, *J. Immunol.* **127**:539-46; Brown et al., 1980, *J. Biol. Chem.* **255**:4980-83; Yeh et al., 1976, *PNAS* **76**:2927-31; and Yeh et al., 1982, *Int. J. Cancer* **29**:269-75), the human B cell hybridoma technique (Kozbor et al., 1983, *Immunol. Today* **4**:72), the EBV-hybridoma technique (Cole et al., 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques.  
20

The technology for producing hybridomas is well-known (see generally R. H. Kenneth, 1980, *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, NY; E.A. Lerner, 1981, *Yale J. Biol. Med.*, **54**:387-402; M.L. Geffer et al., 1977, *Somatic Cell Genet.* **3**:231-  
25 36). In general, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a Gene 216 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds Gene 216 polypeptides or peptides.

30 Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-

Gene 216 monoclonal antibody (see, e.g., G. Galfre et al., 1977, *Nature* 266:55052; Gefter et al., 1977; Lerner, 1981; Kenneth, 1980). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods. Typically, the immortal cell line (e.g., a myeloma cell line) is derived  
5 from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin, and  
10 thymidine (HAT medium). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653, or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC (American Type Culture Collection, Manassas, VA). Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes  
15 using polyethylene glycol (PEG). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for  
20 antibodies that bind Gene 216 polypeptides or peptides, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-Gene 216 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage  
25 display library) with Gene 216 to thereby isolate immunoglobulin library members that bind Gene 216. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; and the Stratagene SurfZAP™ Phage Display Kit, Catalog No. 240612).

30 Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in,

for example, Ladner et al. U.S. Pat. No. 5,223,409; Kang et al. PCT International Publication No. WO 92/18619; Dower et al. PCT International Publication No. WO 91/17271; Winter et al. PCT International Publication WO 92/20791; Markland et al. PCT International Publication No. WO 92/15679;  
5 Breitling et al. PCT International Publication WO 93/01288; McCafferty et al. PCT International Publication No. WO 92/01047; Garrard et al. PCT International Publication No. WO 92/09690; Ladner et al. PCT International Publication No. WO 90/02809; Fuchs et al., 1991, *Bio/Technology* **9**:1370-1372; Hay et al., 1992, *Hum. Antibod. Hybridomas* **3**:81-85; Huse et al., 1989,  
10 *Science* **246**:1275-1281; Griffiths et al., 1993, *EMBO J* **12**:725-734; Hawkins et al., 1992, *J. Mol. Biol.* **226**:889-896; Clarkson et al., 1991, *Nature* **352**:624-628; Gram et al., 1992, *PNAS* **89**:3576-3580; Garrad et al., 1991, *Bio/Technology* **9**:1373-1377; Hoogenboom et al., 1991, *Nuc. Acid Res.* **19**:4133-4137; Barbas et al., 1991, *PNAS* **88**:7978-7982; and McCafferty et al.,  
15 1990, *Nature* **348**:552-55.

Additionally, recombinant anti-Gene 216 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and  
20 humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al. International Application No. PCT/US86/02269; Akira, et al. European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al. European Patent Application 173,494; Neuberger et  
25 al. PCT International Publication No. WO 86/01533; Cabilly et al. U.S. Pat. No. 4,816,567; Cabilly et al. European Patent Application 125,023; Better et al., 1988, *Science* **240**:1041-1043; Liu et al., 1987, *PNAS* **84**:3439-3443; Liu et al., 1987, *J. Immunol.* **139**:3521-3526; Sun et al., 1987, *PNAS* **84**:214-218; Nishimura et al., 1987, *Canc. Res.* **47**:999-1005; Wood et al., 1985, *Nature*  
30 **314**:446-449; and Shaw et al., 1988, *J. Natl. Cancer Inst.* **80**:1553-1559; S.L. Morrison, 1985, *Science* **229**:1202-1207; Oi et al., 1986, *BioTechniques* **4**:214;

Winter U.S. Pat. No. 5,225,539; Jones et al., 1986, *Nature* **321**:552-525; Verhoeyan et al., 1988, *Science* **239**:1534; and Bcidler et al., 1988, *J. Immunol.* **141**:4053-4060.

An anti-Gene 216 antibody (e.g., monoclonal antibody) can be used to isolate Gene 216 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-Gene 216 antibody can also facilitate the purification of natural Gene 216 polypeptide from cells and of recombinantly produced Gene 216 polypeptides or peptides expressed in host cells. Further, an anti-Gene 216 antibody can be used to detect Gene 216 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the Gene 216 protein. Anti-Gene 216 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen as described in detail herein. In addition, and anti-Gene 216 antibody can be used as therapeutics for the treatment of diseases related to abnormal Gene 216 expression or function, e.g., asthma.

### Ligands

The Gene 216 polypeptides, polynucleotides, variants, or fragments thereof, can be used to screen for ligands (e.g., agonists, antagonists, or inhibitors) that modulate the levels or activity of the Gene 216 polypeptide. In addition, these Gene 216 molecules can be used to identify endogenous ligands that bind to Gene 216 polypeptides or polynucleotides in the cell. In one aspect of the present invention, the full-length Gene 216 polypeptide (e.g., SEQ ID NO:4) is used to identify ligands. Alternatively, variants or fragments of a Gene 216 polypeptide are used. Such fragments may comprise, for example, one or more domains of the Gene 216 polypeptide (e.g., the pre-, pro-, catalytic, cysteine-rich, disintegrin, EGF, transmembrane, and cytoplasmic domains) disclosed herein. Of particular interest are screening assays that identify agents that have relatively low levels of toxicity in human cells. A wide variety of assays may be used for this purpose, including *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, and the



like.

The term "ligand" as used herein describes any molecule, protein, peptide, or compound with the capability of directly or indirectly altering the physiological function, stability, or levels of the Gene 216 polypeptide. Ligands that bind to the Gene 216 polypeptides or polynucleotides of the invention are potentially useful in diagnostic applications and/or pharmaceutical compositions, as described in detail herein. Ligands may encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Such ligands can comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. Ligands often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Ligands can also comprise biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs, or combinations thereof.

Ligands may include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam et al., 1991, *Nature* **354**:82-84; Houghten et al., 1991, *Nature* **354**:84-86) and combinatorial chemistry-derived molecular libraries made of D- and/or L-configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang et al, 1993, *Cell* **72**:767-778); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, F(ab')<sub>2</sub>, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules.

Ligands can be obtained from a wide variety of sources including libraries of synthetic or natural compounds. Synthetic compound libraries are commercially available from, for example, Maybridge Chemical Co. (Trevillet,

Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich Chemical Company, Inc. (Milwaukee, WI). Natural compound libraries comprising bacterial, fungal, plant or animal extracts are available from, for example, Pan Laboratories (Bothell, WA). In addition, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides.

Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts can be readily produced. Methods for the synthesis of molecular libraries are readily available (see, e.g., DeWitt et al., 1993, *Proc. Natl. Acad. Sci. USA* **90**:6909; Erb et al., 1994, *Proc. Natl. Acad. Sci. USA* **91**:11422; Zuckermann et al., 1994, *J. Med. Chem.* **37**:2678; Cho et al., 1993, *Science* **261**:1303; Carell et al., 1994, *Angew. Chem. Int. Ed. Engl.* **33**:2059; Carell et al., 1994, *Angew. Chem. Int. Ed. Engl.* **33**:2061; and in Gallop et al., 1994, *J. Med. Chem.* **37**:1233). In addition, natural or synthetic compound libraries and compounds can be readily modified through conventional chemical, physical and biochemical means (see, e.g., Blondelle et al., 1996, *Trends in Biotech.* **14**:60), and may be used to produce combinatorial libraries. In another approach, previously identified pharmacological agents can be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, and the analogs can be screened for Gene 216-modulating activity.

Numerous methods for producing combinatorial libraries are known in the art, including those involving biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds (K. S. Lam, 1997, *Anticancer Drug Des.*

12:145).

Libraries may be screened in solution (e.g., Houghten, 1992, *Biotechniques* **13**:412-421), or on beads (Lam, 1991, *Nature* **354**:82-84), chips (Fodor, 1993, *Nature* **364**:555-556), bacteria or spores (Ladner U.S. Pat. No. 5,223,409), plasmids (Cull et al., 1992, *Proc. Natl. Acad. Sci. USA* **89**:1865-1869), or on phage (Scott and Smith, 1990, *Science* **249**:386-390; Devlin, 1990, *Science* **249**:404-406; Cwirla et al., 1990, *Proc. Natl. Acad. Sci. USA* **97**:6378-6382; Felici, 1991, *J. Mol. Biol.* **222**:301-310; Ladner, *supra*).

Where the screening assay is a binding assay, a Gene 216 polypeptide, polynucleotide, analog, or fragment thereof, may be joined to a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, e.g. magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin, etc. For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in accordance with known procedures.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc., that are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The components are added in any order that produces the requisite binding. Incubations are performed at any temperature that facilitates optimal activity, typically between 4° and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high-throughput screening. Normally, between 0.1 and 1 hr will be sufficient. In general, a plurality of assay mixtures is run in parallel with different agent concentrations to obtain a differential response to these concentrations. Typically, one of these concentrations serves as a negative control, i.e. at zero concentration or below the level of detection.

To perform cell-free ligand screening assays, it may be desirable to immobilize either the Gene 216 polypeptide, polynucleotide, or fragment to a surface to facilitate identification of ligands that bind to these molecules, as well as to accommodate automation of the assay. For example, a fusion protein comprising a Gene 216 polypeptide and an affinity tag can be produced. In one embodiment, a glutathione-S-transferase/phosphodiesterase fusion protein comprising a Gene 216 polypeptide is adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione-derivatized microtiter plates. Cell lysates (e.g., containing <sup>35</sup>S-labeled polypeptides) are added to the Gene 216-coated beads under conditions to allow complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the Gene 216-coated beads are washed to remove any unbound polypeptides, and the amount of immobilized radiolabel is determined. Alternatively, the complex is dissociated and the radiolabel present in the supernatant is determined. In another approach, the beads are analyzed by SDS-PAGE to identify Gene 216-binding polypeptides.

Ligand-binding assays can be used to identify agonist or antagonists that alter the function or levels of the Gene 216 polypeptide. Such assays are designed to detect the interaction of test agents with Gene 216 polypeptides, polynucleotides, analogs, or fragments thereof. Interactions may be detected by direct measurement of binding. Alternatively, interactions may be detected by indirect indicators of binding, such as stabilization/destabilization of protein structure, or activation/inhibition of biological function. Non-limiting examples of useful ligand-binding assays are detailed below.

Ligands that bind to Gene 216 polypeptides, polynucleotides, analogs, or fragments thereof, can be identified using real-time Bimolecular Interaction Analysis (BIA; Sjolander et al., 1991, *Anal. Chem.* **63**:2338-2345; Szabo et al., 1995, *Curr. Opin. Struct. Biol.* **5**:699-705). BIA-based technology (e.g., BIAcore<sup>TM</sup>; LKB Pharmacia, Sweden) allows study of biospecific interactions in real time, without labeling. In BIA, changes in the optical phenomenon surface plasmon resonance (SPR) is used determine real-time interactions of

biological molecules.

Ligands can also be identified by scintillation proximity assays (SPA, described in U.S. Patent No. 4,568,649). In a modification of this assay that is currently undergoing development, chaperonins are used to distinguish  
5 folded and unfolded proteins. A tagged protein is attached to SPA beads, and test agents are added. The bead is then subjected to mild denaturing conditions (such as, e.g., heat, exposure to SDS, etc.) and a purified labeled chaperonin is added. If a test agent binds to a target, the labeled chaperonin will not bind; conversely, if no test agent binds, the protein will undergo some  
10 degree of denaturation and the chaperonin will bind.

Ligands can also be identified using a binding assay based on mitochondrial targeting signals (Hurt et al., 1985, *EMBO J.* 4:2061-2068; Eilers and Schatz, 1986, *Nature* 322:228-231). In a mitochondrial import assay, expression vectors are constructed in which nucleic acids encoding particular  
15 target proteins are inserted downstream of sequences encoding mitochondrial import signals. The chimeric proteins are synthesized and tested for their ability to be imported into isolated mitochondria in the absence and presence of test compounds. A test compound that binds to the target protein should inhibit its uptake into isolated mitochondria *in vitro*.

20 The ligand-binding assay described in Fodor et al., 1991, *Science* 251:767-773, which involves testing the binding affinity of test compounds for a plurality of defined polymers synthesized on a solid substrate, can also be used.

Ligands that bind to Gene 216 polypeptides or peptides can be  
25 identified using two-hybrid assays (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al., 1993, *Cell* 72:223-232; Madura et al., 1993, *J. Biol. Chem.* 268:12046-12054; Bartel et al., 1993, *Biotechniques* 14:920-924; Iwabuchi et al., 1993, *Oncogene* 8:1693-1696; and Brent WO 94/10300). The two-hybrid system relies on the reconstitution of transcription activation activity by association of  
30 the DNA-binding and transcription activation domains of a transcriptional activator through protein-protein interaction. The yeast GAL4 transcriptional

activator may be used in this way, although other transcription factors have been used and are well known in the art. To carryout the two-hybrid assay, the GAL4 DNA-binding domain, and the GAL4 transcription activation domain are expressed, separately, as fusions to potential interacting polypeptides.

5           In one embodiment, the "bait" protein comprises a Gene 216 polypeptide fused to the GAL4 DNA-binding domain. The "fish" protein comprises, for example, a human cDNA library encoded polypeptide fused to the GAL4 transcription activation domain. If the two, coexpressed fusion proteins interact in the nucleus of a host cell, a reporter gene (e.g. LacZ) is  
10           activated to produce a detectable phenotype. The host cells that show two-hybrid interactions can be used to isolate the containing plasmids containing the cDNA library sequences. These plasmids can be analyzed to determine the nucleic acid sequence and predicted polypeptide sequence of the candidate ligand. Alternatively, methods such as the three-hybrid (Licitra et al.,  
15           1996, *Proc. Natl. Acad. Sci. USA* **93**:12817-12821), and reverse two-hybrid (Vidal et al., 1996, *Proc. Natl. Acad. Sci. USA* **93**:10315-10320) systems may be used. Commercially available two-hybrid systems such as the CLONTECH Matchmaker™ systems and protocols (CLONTECH Laboratories, Inc., Palo Alto, CA) may be also be used (see also, A.R. Mendelsohn et al., 1994, *Curr.*  
20           *Op. Biotech.* **5**:482; E.M. Phizicky et al., 1995, *Microbiological Rev.* **59**:94; M. Yang et al., 1995, *Nucleic Acids Res.* **23**:1152; S. Fields et al., 1994, *Trends Genet.* **10**:286; and U.S. Patent No. 6,283,173 and 5,468,614).

          Several methods of automated assays have been developed in recent years so as to permit screening of tens of thousands of test agents in a short  
25           period of time. High-throughput screening methods are particularly preferred for use with the present invention. The ligand-binding assays described herein can be adapted for high-throughput screens, or alternative screens may be employed. For example, continuous format high throughput screens (CF-HTS) using at least one porous matrix allows the researcher to test large numbers  
30           of test agents for a wide range of biological or biochemical activity (see United States Patent No. 5,976,813 to Beutel et al.). Moreover, CF-HTS can be used

to perform multi-step assays.

### **Diagnostics**

As discussed herein, chromosomal region 20p13-p12 has been genetically linked to a variety of diseases and disorders, including asthma.

- 5 The present invention provides nucleic acids and antibodies that can be useful in diagnosing individuals with aberrant Gene 216 expression. In particular, the disclosed SNPs can be used to diagnose chromosomal abnormalities linked to these diseases.

- Antibody-based diagnostic methods: In a further embodiment of the present invention, antibodies which specifically bind to the Gene 216 polypeptide may be used for the diagnosis of conditions or diseases characterized by underexpression or overexpression of the Gene 216 polynucleotide or polypeptide, or in assays to monitor patients being treated with a Gene 216 polypeptide or peptide, or a Gene 216 agonist, antagonist, or inhibitor.

The antibodies useful for diagnostic purposes may be prepared in the same manner as those for use in therapeutic methods, described herein.

- Antibodies may be raised to the full-length Gene 216 polypeptide sequence (e.g., SEQ ID NO:4). Alternatively, the antibodies may be raised to fragments or variants of the Gene 216 polypeptide. In one aspect of the invention, antibodies are prepared to bind to a Gene 216 polypeptide fragment comprising one or more domains of the Gene 216 polypeptide (e.g., pre-, pro-, catalytic, disintegrin, cysteine-rich, EGF, transmembrane, and cytoplasmic domains) described herein.

- 25 Diagnostic assays for the Gene 216 polypeptide include methods that utilize the antibody and a label to detect the protein in biological samples (e.g., human body fluids, cells, tissues, or extracts of cells or tissues). The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule.

- 30 A wide variety of reporter molecules that are known in the art may be used, several of which are described herein.

The invention provides methods for detecting disease-associated antigenic components in a biological sample, which methods comprise the steps of: 1) contacting a sample suspected to contain a disease-associated antigenic component with an antibody specific for an disease-associated antigen, extracellular or intracellular, under conditions in which an antigen-antibody complex can form between the antibody and disease-associated antigenic components in the sample; and 2) detecting any antigen-antibody complex formed in step (1) using any suitable means known in the art, wherein the detection of a complex indicates the presence of disease-associated antigenic components in the sample. It will be understood that assays that utilize antibodies directed against altered Gene 216 amino acid sequences (i.e., epitopes encoded by SNPs, mutations, or variants) are within the scope of the invention.

Many immunoassay formats are known in the art, and the particular format used is determined by the desired application. An immunoassay can use, for example, a monoclonal antibody directed against a single disease-associated epitope, a combination of monoclonal antibodies directed against different epitopes of a single disease-associated antigenic component, monoclonal antibodies directed towards epitopes of different disease-associated antigens, polyclonal antibodies directed towards the same disease-associated antigen, or polyclonal antibodies directed towards different disease-associated antigens. Protocols can also, for example, use solid supports, or may involve immunoprecipitation.

In accordance with the present invention, "competitive" (U.S. Pat. Nos. 3,654,090 and 3,850,752), "sandwich" (U.S. Pat. No. 4,016,043), and "double antibody," or "DASP" assays may be used. Several procedures for measuring the Gene 216 polypeptide (e.g., ELISA, RIA, and FACS) are known in the art and provide a basis for diagnosing altered or abnormal levels of Gene 216 polypeptide expression. Normal or standard values for Gene 216 polypeptide expression are established by incubating biological samples taken from normal subjects, preferably human, with antibody to the Gene polypeptide under



conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods; photometric means are preferred. Levels of the Gene 216 polypeptide expressed in the subject sample, negative control (normal) sample, and positive control (disease) sample are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

Typically, immunoassays use either a labeled antibody or a labeled antigenic component (e.g., that competes with the antigen in the sample for binding to the antibody). A number of fluorescent materials are known and can be utilized as labels for antibodies or polypeptides. These include, for example, Cy3, Cy5, Alexa, BODIPY, fluorescein (e.g., FluorX, DTAF, and FITC), rhodamine (e.g., TRITC), auramine, Texas Red, AMCA blue, and Lucifer Yellow. Antibodies or polypeptides can also be labeled with a radioactive element or with an enzyme. Preferred isotopes include  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{36}\text{Cl}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{90}\text{Y}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ , and  $^{186}\text{Re}$ . Preferred enzymes include peroxidase,  $\beta$ -glucuronidase,  $\beta$ -D-glucosidase,  $\beta$ -D-galactosidase, urease, glucose oxidase plus peroxidase, and alkaline phosphatase (see, e.g., U.S. Pat. Nos. 3,654,090; 3,850,752 and 4,016,043). Enzymes can be conjugated by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde, and the like. Enzyme labels can be detected visually, or measured by calorimetric, spectrophotometric, fluorospectrophotometric, amperometric, or gasometric techniques. Other labeling systems, such as avidin/biotin, Tyramide Signal Amplification (TSA<sup>TM</sup>), are known in the art, and are commercially available (see, e.g., ABC kit, Vector Laboratories, Inc., Burlingame, CA; NEN<sup>®</sup> Life Science Products, Inc., Boston, MA).

Kits suitable for antibody-based diagnostic applications typically include one or more of the following components:

(1) Antibodies: The antibodies may be pre-labeled; alternatively, the antibody may be unlabeled and the ingredients for labeling may be included in the kit in separate containers, or a secondary, labeled antibody is

provided; and

(2) Reaction components: The kit may also contain other suitably packaged reagents and materials needed for the particular immunoassay protocol, including solid-phase matrices, if applicable, and standards.

5           The kits referred to above may include instructions for conducting the test. Furthermore, in preferred embodiments, the diagnostic kits are adaptable to high-throughput and/or automated operation.

Nucleic-acid-based diagnostic methods: The invention provides methods for altered levels or sequences of Gene 216 nucleic acids in a  
10       sample, such as in a biological sample, which methods comprise the steps of: 1) contacting a sample suspected to contain a disease-associated nucleic acid with one or more disease-associated nucleic acid probes under conditions in which hybrids can form between any of the probes and disease-associated nucleic acid in the sample; and 2) detecting any hybrids  
15       formed in step (1) using any suitable means known in the art, wherein the detection of hybrids indicates the presence of the disease-associated nucleic acid in the sample. To detect disease-associated nucleic acids present in low levels in biological samples, it may be necessary to amplify the disease-associated sequences or the hybridization signal as part of the  
20       diagnostic assay. Techniques for amplification are known to those of skill in the art.

          The presence of Gene 216 polynucleotide sequences can be detected by DNA-DNA or DNA-RNA hybridization, or by amplification using probes or primers comprising at least a portion of a Gene 216  
25       polynucleotide, or a sequence complementary thereto. In particular, nucleic acid amplification-based assays can use Gene 216 oligonucleotides or oligomers to detect transformants containing Gene 216 DNA or RNA. Gene 216 nucleic acids useful as probes in diagnostic methods include oligonucleotides at least 15 nucleotides in length, preferably at least 20  
30       nucleotides in length, and most preferably at least 25-55 nucleotides in length, that hybridize specifically with Gene 216 nucleic acids.

Several methods can be used to produce specific probes for Gene 216 polynucleotides. For example, labeled probes can be produced by oligo-labeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, Gene 216 polynucleotide sequences (e.g., SEQ ID NO:1 or SEQ ID NO:6), or any portions or fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase, such as T7, T3, or SP(6) and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (e.g., from Amersham-Pharmacia; Promega Corp.; and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels which may be used include radionucleotides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

A sample to be analyzed, such as, for example, a tissue sample (e.g., hair or buccal cavity) or body fluid sample (e.g., blood or saliva), may be contacted directly with the nucleic acid probes. Alternatively, the sample may be treated to extract the nucleic acids contained therein. It will be understood that the particular method used to extract DNA will depend on the nature of the biological sample. The resulting nucleic acid from the sample may be subjected to gel electrophoresis or other size separation techniques, or, the nucleic acid sample may be immobilized on an appropriate solid matrix without size separation.

Kits suitable for nucleic acid-based diagnostic applications typically include the following components:

- (1) Probe DNA: The probe DNA may be prelabeled; alternatively, the probe DNA may be unlabeled and the ingredients for labeling may be included in the kit in separate containers; and
- (2) Hybridization reagents: The kit may also contain other suitably packaged reagents and materials needed for the particular hybridization protocol, including solid-phase matrices, if applicable, and standards.

In cases where a disease condition is suspected to involve an alteration of the Gene 216 nucleotide sequence, specific oligonucleotides may be constructed and used to assess the level of disease mRNA in cells affected or other tissue affected by the disease. For example, PCR can be  
5 used to test whether a person has a disease-related polymorphism (i.e., mutation).

For PCR analysis, Gene 216 oligonucleotides may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably comprise two nucleotide sequences, one  
10 with a sense orientation (5' → 3') and another with an antisense orientation (3' → 5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA  
15 sequences.

In accordance with PCR analysis, two oligonucleotides are synthesized by standard methods or are obtained from a commercial supplier of custom-made oligonucleotides. The length and base composition are determined by standard criteria using the Oligo 4.0 primer  
20 Picking program (W. Rychlik, 1992; available from Molecular Biology Insights, Inc., Cascade, CO). One of the oligonucleotides is designed so that it will hybridize only to the disease gene DNA under the PCR conditions used. The other oligonucleotide is designed to hybridize a segment of genomic DNA such that amplification of DNA using these oligonucleotide  
25 primers produces a conveniently identified DNA fragment. Samples may be obtained from hair follicles, whole blood, or the buccal cavity. The DNA fragment generated by this procedure is sequenced by standard techniques.

In one particular aspect, Gene 216 oligonucleotides can be used to perform Genetic Bit Analysis (GBA) of Gene 216 in accordance with  
30 published methods (T.T. Nikiforov et al., 1994, *Nucleic Acids Res.* **22**(20):4167-75; T.T. Nikiforov et al., 1994, *PCR Methods Appl.* **3**(5):285-

91). In PCR-based GBA, specific fragments of genomic DNA containing the polymorphic site(s) are first amplified by PCR using one unmodified and one phosphorothioate-modified primer. The double-stranded PCR product is rendered single-stranded and then hybridized to immobilized oligonucleotide primer in wells of a multi-well plate. The primer is designed to anneal immediately adjacent to the polymorphic site of interest. The 3' end of the primer is extended using a mixture of individually labeled dideoxynucleoside triphosphates. The label on the extended base is then determined. Preferably, GBA is performed using semi-automated ELISA or biochip formats (see, e.g., S.R. Head et al., 1997, *Nucleic Acids Res.* **25**(24):5065-71; T.T. Nikiforov et al., 1994, *Nucleic Acids Res.* **22**(20):4167-75).

Other amplification techniques besides PCR may be used as alternatives, such as ligation-mediated PCR or techniques involving Q-beta replicase (Cahill et al., 1991, *Clin. Chem.*, **37**(9):1482-5). Products of amplification can be detected by agarose gel electrophoresis, quantitative hybridization, or equivalent techniques for nucleic acid detection known to one skilled in the art of molecular biology (Sambrook et al., 1989). Other alterations in the disease gene may be diagnosed by the same type of amplification-detection procedures, by using oligonucleotides designed to contain and specifically identify those alterations.

Gene 216 polynucleotides may also be used to detect and quantify levels of Gene 216 mRNA in biological samples in which altered expression of Gene 216 polynucleotide may be correlated with disease. These diagnostic assays may be used to distinguish between the absence, presence, increase, and decrease of Gene 216 mRNA levels, and to monitor regulation of Gene 216 polynucleotide levels during therapeutic treatment or intervention. For example, Gene 216 polynucleotide sequences, or fragments, or complementary sequences thereof, can be used in Southern or Northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or biochip assays utilizing fluids or tissues from patient biopsies to detect the status of, e.g., levels or

overexpression of Gene 216, or to detect altered Gene 216 expression. Such qualitative or quantitative methods are well known in the art (G.H. Keller and M.M. Manak, 1993, *DNA Probes*, 2<sup>nd</sup> Ed, Macmillan Publishers Ltd., England; D.W. Dieffenbach and G. S. Dveksler, 1995, *PCR Primer: A Laboratory*  
5 *Manual*, Cold Spring Harbor Press, Plainview, NY; B.D. Hames and S.J. Higgins, 1985, *Gene Probes 1, 2*, IRL Press at Oxford University Press, Oxford, England).

Methods suitable for quantifying the expression of Gene 216 include radiolabeling or biotinylating nucleotides, co-amplification of a control nucleic  
10 acid, and standard curves onto which the experimental results are interpolated (P.C. Melby et al., 1993, *J. Immunol. Methods* **159**:235-244; and C. Duplaa et al., 1993, *Anal. Biochem.* 229-236). The speed of quantifying multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric  
15 or colorimetric response gives rapid quantification.

In accordance with these methods, the specificity of the probe, i.e., whether it is made from a highly specific region (e.g., at least 8 to 10 or 12 or 15 contiguous nucleotides in the 5' regulatory region), or a less specific region (e.g., especially in the 3' coding region), and the stringency of the hybridization  
20 or amplification (e.g., high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding the Gene 216 polypeptide, alleles thereof, or related sequences.

In a particular aspect, a Gene 216 nucleic acid sequence, or a sequence complementary thereto, or fragment thereof, may be useful in  
25 assays that detect Gene 216-related diseases such as asthma. The Gene 216 polynucleotide can be labeled by standard methods, and added to a biological sample from a subject under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample can be washed and the signal is quantified and compared with a standard value. If the  
30 amount of signal in the test sample is significantly altered from that of a comparable negative control (normal) sample, the altered levels of Gene 216

nucleotide sequence can be correlated with the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular prophylactic or therapeutic regimen in animal studies, in clinical trials, or for an individual patient.

5           To provide a basis for the diagnosis of a disease associated with altered expression of Gene 216, a normal or standard profile for expression is established. This may be accomplished by incubating biological samples taken from normal subjects, either animal or human, with a sequence complementary to the Gene 216 polynucleotide, or a fragment thereof, under conditions  
10           suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are  
15           symptomatic for the disease. Deviation between standard and subject (patient) values is used to establish the presence of the condition.

          Once the disease is diagnosed and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is  
20           observed in a normal individual. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

          With respect to diseases such as asthma, the presence of an abnormal amount of Gene 216 transcript in a biological sample (e.g., body fluid, cells,  
25           tissues, or cell or tissue extracts) from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the  
30           development or further progression of the disease.

Microarrays: In another embodiment of the present invention, oligonucleotides, or longer fragments derived from the Gene 216 polynucleotide sequence described herein may be used as targets in a microarray (e.g., biochip) system. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disease, to diagnose disease, and to develop and monitor the activities of therapeutic or prophylactic agents. Preparation and use of microarrays have been described in WO 95/11995 to Chee et al.; D.J. Lockhart et al., 1996, *Nature Biotechnology* **14**:1675-1680; M. Schena et al., 1996, *Proc. Natl. Acad. Sci. USA* **93**:10614-10619; U.S. Patent No. 6,015,702 to P. Lal et al; J. Worley et al., 2000, *Microarray Biochip Technology*, M. Schena, ed., Biotechniques Book, Natick, MA, pp. 65-86; Y.H. Rogers et al., 1999, *Anal. Biochem.* **266**(1):23-30; S.R. Head et al., 1999, *Mol. Cell. Probes.* **13**(2):81-7; S.J. Watson et al., 2000, *Biol. Psychiatry* **48**(12):1147-56.

In one application of the present invention, microarrays containing arrays of Gene 216 polynucleotide sequences can be used to measure the expression levels of Gene 216 in an individual. In particular, to diagnose an individual with a Gene 216-related condition or disease, a sample from a human or animal (containing nucleic acids, e.g., mRNA) can be used as a probe on a biochip containing an array of Gene 216 polynucleotides (e.g., DNA) in decreasing concentrations (e.g., 1 ng, 0.1 ng, 0.01 ng, etc.). The test sample can be compared to samples from diseased and normal samples. Biochips can also be used to identify Gene 216 mutations or polymorphisms in a population, including but not limited to, deletions, insertions, and mismatches. For example, mutations can be identified by: 1) placing Gene 216 polynucleotides of this invention onto a biochip; 2) taking a test sample (containing, e.g., mRNA) and adding the sample to the biochip; 3) determining if the test samples hybridize to the Gene 216 polynucleotides attached to the



chip under various hybridization conditions (see, e.g., V.R. Chechetkin et al., 2000, *J. Biomol. Struct. Dyn.* **18**(1):83-101). Alternatively microarray sequencing can be performed (see, e.g., E.P. Diamandis, 2000, *Clin. Chem.* **46**(10):1523-5).

5        Chromosome mapping: In another application of this invention, the Gene 216 nucleic acid sequence, or a complementary sequence, or fragment thereof, can be used as probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to human artificial  
10       chromosome constructions (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial PI constructions, or single chromosome cDNA libraries (see C.M. Price, 1993, *Blood Rev.*, **7**:127-134 and by B.J. Trask, 1991, *Trends Genet.* **7**:149-154).

15       In another of its aspects, the invention relates to a diagnostic kit for detecting Gene 216 polynucleotide or polypeptide as it relates to a disease or susceptibility to a disease, particularly asthma. Also related is a diagnostic kit that can be used to detect or assess asthma conditions. Such kits comprise one or more of the following:

- 20       (a) a Gene 216 polynucleotide, preferably the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:6, or a fragment thereof; or
- (b) a nucleotide sequence complementary to that of (a); or
- (c) a Gene 216 polypeptide, preferably the polypeptide of SEQ ID NO:4, or a fragment thereof; or
- 25       (d) an antibody to a Gene 216 polypeptide, preferably to the polypeptide of SEQ ID NO:4, or an antibody bindable fragment thereof. It will be appreciated that in any such kits, (a), (b), (c), or (d) may comprise a substantial component and that instructions for use can be included. The kits may also contain peripheral reagents such as buffers, stabilizers, etc.

30       The present invention also includes a test kit for genetic screening that can be utilized to identify mutations in Gene 216. By identifying patients with mutated Gene 216 DNA and comparing the mutation to a database that

contains known mutations in Gene 216 and a particular condition or disease, identification and/or confirmation of, a particular condition or disease can be made. Accordingly, such a kit would comprise a PCR-based test that would involve transcribing the patients mRNA with a specific primer, and amplifying  
5 the resulting cDNA using another set of primers. The amplified product would be detectable by gel electrophoresis and could be compared with known standards for Gene 216. Preferably, this kit would utilize a patient's blood, serum, or saliva sample, and the DNA would be extracted using standard techniques. Primers flanking a known mutation would then be used to amplify  
10 a fragment of Gene 216. The amplified piece would then be sequenced to determine the presence of a mutation.

Genomic Screening: The use of polymorphic genetic markers linked to the Gene 216 gene is very useful in predicting susceptibility to the diseases genetically linked to 20p13-p12. Similarly, the identification of polymorphic  
15 genetic markers within the Gene 216 gene will allow the identification of specific allelic variants that are in linkage disequilibrium with other genetic lesions that affect one of the disease states discussed herein including respiratory disorders, obesity, and inflammatory bowel disease. SSCP (see below) allows the identification of polymorphisms within the genomic and  
20 coding region of the disclosed gene. The present invention provides sequences for primers that can be used identify exons that contain SNPs, as well as sequences for primers that can be used to identify the sequence change. This information can be used to identify additional SNPs in accordance with the methods disclosed herein. Suitable methods for genomic  
25 screening have also been described by, e.g., Sheffield et al., 1995, *Genet.*, 4:1837-1844; LeBlanc-Straceski et al., 1994, *Genomics*, 19:341-9; Chen et al., 1995, *Genomics*, 25:1-8. In employing these methods, the disclosed reagents can be used to predict the risk for disease (e.g., respiratory disorders, obesity, and inflammatory bowel disease) in a population or individual.

### 30 Therapeutics

The present invention provides methods of screening for drugs

comprising contacting such an agent with a novel protein of this invention or fragment thereof and assaying 1) for the presence of a complex between the agent and the protein or fragment, or 2) for the presence of a complex between the protein or fragment and a ligand, by methods well known in the art. In such competitive binding assays the novel protein or fragment is typically labeled. Free protein or fragment is separated from that present in a protein:protein complex, and the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to Gene 216 protein or its interference with protein ligand binding, respectively.

10        This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of specifically binding the Gene 216 protein compete with a test compound for binding to the Gene 216 protein or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more  
15        antigenic determinants of a Gene 216 protein.

          The goal of rational drug design is to produce structural analogs of biologically active proteins of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the protein, or which,  
20        e.g., enhance or interfere with the function of a protein *in vivo* (see, e.g., Hodgson, 1991, *Bio/Technology*, 9:19-21). In one approach, one first determines the three-dimensional structure of a protein of interest or, for example, of the Gene 216 receptor or ligand complex, by x-ray crystallography, by computer modeling or most typically, by a combination  
25        of approaches. Less often, useful information regarding the structure of a protein may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson et al., 1990, *Science*, 249:527-533). In addition, peptides (e.g., Gene 216 protein) are analyzed by an alanine scan  
30        (Wells, 1991, *Methods in Enzymol.*, 202:390-411). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity

is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original Gene 216 protein. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which result in, for example, altered Gene 216 protein activity or stability or which act as inhibitors, agonists, antagonists, etc. of Gene 216 protein activity. By virtue of the availability of cloned Gene 216 gene sequences, sufficient amounts of the Gene 216 protein may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge of the Gene 216 polypeptide sequence will guide those employing computer-modeling techniques in place of, or in addition to x-ray crystallography.

In another aspect of the present invention, cells and animals that carry the Gene 216 gene or an analog thereof can be used as model systems to study and test for substances that have potential as therapeutic agents. After a test substance is administered to animals or applied to the cells, the phenotype of the animals/cells can be determined.

In yet another aspect of this invention, antibodies that specifically react with Gene 216 polypeptide or peptides derived therefrom can be used as therapeutics. In particular, anti-Gene 216 antibodies can be used to block the Gene 216 activity. Anti-Gene 216 antibodies or fragments thereof can be formulated as pharmaceutical compositions and administered to a subject. It is noted that antibody-based therapeutics produced from non-human sources

can cause an undesired immune response in human subjects. To minimize this problem, chimeric antibody derivatives can be produced. Chimeric antibodies combine a non-human animal variable region with a human constant region. Chimeric antibodies can be constructed according to methods  
5 known in the art (see Morrison et al., 1985, *Proc. Natl. Acad. Sci. USA*, **81**:6851; Takeda et al., 1985, *Nature* **314**:452; U.S. Patent No. 4,816,567 of Cabilly et al.; U.S. Patent No. 4,816,397 of Boss et al.; European Patent Publication EP 171496; EP 0173494; United Kingdom Patent GB 2177096B).  
In addition, antibodies can be further "humanized" by any of the techniques  
10 known in the art, (e.g., Teng et al., 1983, *Proc. Natl. Acad. Sci. USA* **80**:7308-7312; Kozbor et al., 1983, *Immunology Today* **4**: 7279; Olsson et al., 1982, *Meth. Enzymol.* **92**:3-16; International Patent Application WO92/06193; EP 0239400). Humanized antibodies can also be obtained from commercial sources (e.g., Scotgen Limited, Middlesex, Great Britain). Immunotherapy with  
15 a humanized antibody may result in increased long-term effectiveness for the treatment of chronic disease situations or situations requiring repeated antibody treatments.

In one embodiment, compositions (e.g., pharmaceutical compositions) for use with the present invention comprise metalloprotease inhibitors, or  
20 analogs or derivatives thereof. Non-limiting examples of metalloprotease inhibitors include: 1) naturally occurring inhibitors, e.g., oprin (J.J. Catanese and L.F. Kress, 1992, *Biochemistry* **31**:410-418; HSF (Y. Yamakawa and T. Omori-Satoh, 1992, *J. Biochem.* **112**:583-589); erinacin (D. Mebs et al., 1996, *Toxicon* **34**:1313-1316; Omori-Satoh et al., 2000, *Toxicon* **38**:1561-1580);  
25 DM40 and DM43 (A.G. Neves-Ferreira et al., 2000, *Biochem. Biophys. Acta.* **1473**:309-320); citrate (B. Francis et al., 1992, *Toxicon* **30**:1239-1246); TIMP-1 and TIMP-2 (R.V. Ward et al., 1991, *Biochem J.* **278**, Pt 1:179-873); pyrophosphate (G.S. Makowski and M.L. Ramsby, 1999, *Inflammation* **23**:333-360); proglutamyl peptides such as pyroGlu-Asn-Trp-OH and pyroGlu-Glu-Trp-  
30 OH (A. Robeva et al., 1991, *Biomed. Biochem. Acta.* **50**:769-773); 2) peptide analogs and derivatives, e.g., 2-distereomeric furan-2-carbonylamino-3-

oxohexahydroindolizino[8,7-b]indole carboxylates (S. D'Alessio et al., 2001, *Eur. J. Med. Chem.* **36**:43-53); phosphonate and carboxylate derivatives of pyroGlu-Asn-Trp-OH (D'Alessio et al., 2001); POL 647 and POL 656 (F.X. Gomis-Ruth et al., 1998, *Prot. Sci.* **7**:283-292); cysteine-switches (K. Nomura and N. Suzuki, 1993, *FEBS Lett.* **321**:84-88); 3) hydroxamate compounds, e.g., batimastat/BB-94 (see, e.g., G.F. Beattie et al., 1998, *Clin. Cancer Res.* **8**:1899-1902); prinomastat/AG3340 (see, e.g., R. Scatena, 2000, *Expert Opin. Investig. Drugs* **9**:2159-2165); and 4) other inhibitors, e.g., ortho-substituted macrocyclic lactams (G.M. Ksander, 1997, *J. Med. Chem.* **40**:495-505); diketopiperazine (DKP) (A.K. Szardenings et al., 1998, *J. Med. Chem.* **41**(13):2194-200; alendronate/PCP (Makowski and Ramsby, 1999); and CT1746 (Z. An et al., 1997, *Clin. Exp. Metastasis* **15**:184-195).

In particular, the determined structures of metalloproteases and metalloprotease inhibitors can be used to devise Gene 216-targeted inhibitors (i.e., by rational drug design; see Szardenings et al, 1998). Structural information can be found in, e.g., C. Oefner et al., 2000, *J. Mol. Biol.* **296**(2):341-9; B. Wu et al., 2000, *J. Mol. Biol.* **295**(2):257-68; L. Chen et al., 1999, *J. Mol. Biol.* **293**(3):545-57; C. Fernandez-Catalan et al., 1998, *EMBO J.* **17**(17):5238-48; S. Arumugam et al., 1998, *Biochemistry* **37**(27):9650-7; Gohlke et al., 1996, *FEBS Lett.* **378**:126-130; Gomis-Ruth et al., 1998; F.X. Gomis-Ruth et al, 1993, *EMBO J.* **12**:4151-4157; F.X. Gomis-Ruth et al, 1996, *J. Mol. Biol.* **264**:556-566; K. Maskos et al., 1998, *Proc. Natl. Acad. Sci. USA* **95**(7):3408-12; F.X. Gomis-Ruth et al, 1997, *Nature* **389**:77-80; M. Betz et al., 1997, *Eur. J. Biochem.* **247**(1):356-63; B. Lovejoy et al., 1994, *Biochemistry* **33**(27):8207-17. Structures of zinc metalloproteases are also found in Molecular Modeling DataBase (MMDB) at the NCBI web site <http://www.ncbi.nlm.nih.gov:80/Structure/MMDB/mmdb.shtml> (e.g. Accession Nos. 1D5J, 1D8F, 1D7X, 1BSK, 2TLX, 1TLX, 1BUD, 1BSW, 1UEA, 4AIG, 3AIG, 2AIG, 1KUH, 1DTH, 1UMS, 1UMT, 7TLN, 6TMN, 5TMN, 5TLN, 4TMN, 4TLN, 3TMN, 2TMN, 1TMN, 1TLP, 1IAG, 1HYT, 1AST, 8TLN, 1THL). In an alternative approach, the binding specificity of TIMP proteins can be

engineered to produce inhibitors that specifically inactivate Gene 216 polypeptide (see, e.g., H. Nagase et al., 1999, *Ann. NY Acad. Sci.* **878**:1-11; G.S. Butler et al., 1999, *J. Biol. Chem.* **274**(29):20391-20396).

In another embodiment of the present invention, compositions (e.g., pharmaceutical compositions) for use with the present invention comprise  
5 disintegrin agonists, or analogs or derivatives thereof. The determined structures of disintegrin proteins and domains can be used to devise Gene 216 disintegrin-targeted agonists (i.e., by rational drug design). Such structural information can be found in R.A. Atkinson et al., 1994, *Int. J. Pept. Protein Res.*  
10 **43**:563-72; V. Saudek et al., 1991, *Eur. J. Biochem.* **202**:329-38; H. Minoux et al., 2000, *J. Comput. Aided Mol. Des.* **14**:317-27.

The present invention contemplates compositions comprising a Gene 216 polynucleotide, polypeptide, antibody, ligand (e.g., agonist, antagonist, or inhibitor), or fragments, variants, or analogs thereof, and a physiologically  
15 acceptable carrier, excipient, or diluent as described in detail herein. The present invention further contemplates pharmaceutical compositions useful in practicing the therapeutic methods of this invention. Preferably, a pharmaceutical composition includes, in admixture, a pharmaceutically acceptable excipient (carrier) and one or more of a Gene 216 polypeptide,  
20 polynucleotide, ligand, antibody, or fragment or variant thereof, as described herein, as an active ingredient. The preparation of pharmaceutical compositions that contain Gene 216-related reagents as active ingredients is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions, however, solid forms  
25 suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified. The active therapeutic ingredient is often mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof.  
30 In addition, if desired, the composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH-buffering agents, which

enhance the effectiveness of the active ingredient.

A Gene 216 polypeptide, polynucleotide, ligand, antibody, or variant or fragment thereof can be formulated into the pharmaceutical composition as neutralized physiologically acceptable salt forms. Suitable salts include the acid addition salts (i.e., formed with the free amino groups of the polypeptide or antibody molecule) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed from the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The pharmaceutical compositions can be administered systemically by oral or parenteral routes. Non-limiting parenteral routes of administration include subcutaneous, intramuscular, intraperitoneal, intravenous, transdermal, inhalation, intranasal, intra-arterial, intrathecal, enteral, sublingual, or rectal. Intravenous administration, for example, can be performed by injection of a unit dose. The term "unit dose" when used in reference to a pharmaceutical composition of the present invention refers to physically discrete units suitable as unitary dosage for humans, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle.

In one particular embodiment of the present invention, the disclosed pharmaceutical compositions are administered via mucoactive aerosol therapy (see, e.g., M. Fuloria and B.K. Rubin, 2000, *Respir. Care* **45**:868-873; I. Gonda, 2000, *J. Pharm. Sci.* **89**:940-945; R. Dhand, 2000, *Curr. Opin. Pulm. Med.* **6**(1):59-70; B.K. Rubin, 2000, *Respir. Care* **45**(6):684-94; S. Suarez and A.J. Hickey, 2000, *Respir. Care* **45**(6):652-66).

Pharmaceutical compositions are administered in a manner compatible with the dosage formulation, and in a therapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of



the subject's immune system to utilize the active ingredient, and degree of modulation of Gene 216 activity desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and are specific for each individual. However, suitable dosages may range from about 5 0.1 to 20, preferably about 0.5 to about 10, and more preferably one to several, milligrams of active ingredient per kilogram body weight of individual per day and depend on the route of administration. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by repeated doses at one or more hour intervals by a 10 subsequent injection or other administration. Alternatively, continuous intravenous infusions sufficient to maintain concentrations of 10 nM to 10  $\mu$ M in the blood are contemplated. An exemplary pharmaceutical formulation comprises: Gene 216 antagonist or inhibitor (5.0 mg/ml); sodium bisulfite USP (3.2 mg/ml); disodium edetate USP (0.1 mg/ml); and water for injection q.s.a.d. 15 (1.0 ml). As used herein, "pg" means picogram, "ng" means nanogram, " $\mu$ g" means microgram, "mg" means milligram, " $\mu$ l" means microliter, "ml" means milliliter, and "l" means L.

For further guidance in preparing pharmaceutical formulations, see, e.g., Gilman et al. (eds), 1990, *Goodman and Gilman's: The Pharmacological* 20 *Basis of Therapeutics*, 8th ed., Pergamon Press; and *Remington's Pharmaceutical Sciences*, 17th ed., 1990, Mack Publishing Co., Easton, PA; Avis et al. (eds), 1993, *Pharmaceutical Dosage Forms: Parenteral Medications*, Dekker, New York; Lieberman et al. (eds), 1990, *Pharmaceutical Dosage Forms: Disperse Systems*, Dekker, New York.

25       Pharmacogenetics: The Gene 216 polypeptides and polynucleotides are also useful in pharmacogenetic analysis (i.e., the study of the relationship between an individual's genotype and that individual's response to a therapeutic composition or drug). See, e.g., M. Eichelbaum, 1996, *Clin. Exp. Pharmacol. Physiol.* **23**(10-11):983-985, and M.W. Linder, 1997, *Clin. Chem.* 30 **43**(2):254-266. The genotype of the individual can determine the way a therapeutic acts on the body or the way the body metabolizes the therapeutic.

Further, the activity of drug metabolizing enzymes affects both the intensity and duration of therapeutic activity. Differences in the activity or metabolism of therapeutics can lead to severe toxicity or therapeutic failure. Accordingly, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenetic studies in determining whether to administer a Gene 216 polypeptide, polynucleotide, analog, antagonist, inhibitor, or modulator, as well as tailoring the dosage and/or therapeutic or prophylactic treatment regimen.

In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions can be due to a single factor that alters the way the drug act on the body (altered drug action), or a factor that alters the way the body metabolizes the drug (altered drug metabolism). These conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy which results in haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. The gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response. This has been demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme, ultra-rapid metabolizers fail to respond to standard doses. Recent

studies have determined that ultra-rapid metabolism is attributable to CYP2D6 gene amplification.

By analogy, genetic polymorphism or mutation may lead to allelic variants of Gene 216 in the population which have different levels of activity.

5 The Gene 216 polypeptides or polynucleotides thereby allow a clinician to ascertain a genetic predisposition that can affect treatment modality. In addition, genetic mutation or variants at other genes may potentiate or diminish the activity of Gene 216-targeted drugs. Thus, in a Gene 216-based treatment, polymorphism or mutation may give rise to individuals that are more or less  
10 responsive to treatment. Accordingly, dosage would necessarily be modified to maximize the therapeutic effect within a given population containing the polymorphism. As an alternative to genotyping, specific polymorphic polypeptides or polynucleotides can be identified.

To identify genes that modify Gene 216-targeted drug response, several  
15 pharmacogenetic methods can be used. One pharmacogenomics approach, "genome-wide association", relies primarily on a high-resolution map of the human genome. This high-resolution map shows previously identified gene-related markers (e.g., a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which  
20 has two variants). A high-resolution genetic map can then be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, a high-resolution map can be generated from a combination of some ten million  
25 known single nucleotide polymorphisms (SNPs) in the human genome. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In this way, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be  
30 common among such genetically similar individuals (see, e.g., D.R. Pfof et al., 2000, *Trends Biotechnol.* **18**(8):334-8).

As another example, the "candidate gene approach", can be used. According to this method, if a gene that encodes a drug target is known, all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As yet another example, a "gene expression profiling approach", can be used. This method involves testing the gene expression of an animal treated with a drug (e.g., a Gene 216 polypeptide, polynucleotide, analog, or modulator) to determine whether gene pathways related to toxicity have been turned on.

Information obtained from one of the approaches described herein can be used to establish a pharmacogenetic profile, which can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. A pharmacogenetic profile, when applied to dosing or drug selection, can be used to avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a Gene 216 polypeptide, polynucleotide, analog, antagonist, inhibitor, or modulator.

Gene 216 polypeptides or polynucleotides are also useful for monitoring therapeutic effects during clinical trials and other treatment. Thus, the therapeutic effectiveness of an agent that is designed to increase or decrease gene expression, polypeptide levels, or activity can be monitored over the course of treatment using the Gene 216 compositions or modulators. For example, monitoring can be performed by: 1) obtaining a pre-administration sample from a subject prior to administration of the agent; 2) detecting the level of expression or activity of the protein in the pre-administration sample; 3) obtaining one or more post-administration samples from the subject; 4) detecting the level of expression or activity of the polypeptide in the post-administration samples; 5) comparing the level of expression or activity of the polypeptide in the pre-administration sample with the polypeptide in the post-administration sample or samples; and 6) increasing or decreasing the

administration of the agent to the subject accordingly.

Gene Therapy: In recent years, significant technological advances have been made in the area of gene therapy for both genetic and acquired diseases (Kay et al., 1997, *Proc. Natl. Acad. Sci. USA*, **94**:12744-12746). Gene therapy can be defined as the transfer of DNA for therapeutic purposes. Improvement in gene transfer methods has allowed for development of gene therapy protocols for the treatment of diverse types of diseases. Gene therapy has also taken advantage of recent advances in the identification of new therapeutic genes, improvement in both viral and non-viral gene delivery systems, better understanding of gene regulation, and improvement in cell isolation and transplantation. Gene therapy would be carried out according to generally accepted methods as described by, for example, Friedman, 1991, *Therapy for Genetic Diseases*, Friedman, Ed., Oxford University Press, pages 105-121.

Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for introducing DNA into cells such as electroporation, calcium phosphate co-precipitation, and viral transduction are known in the art, and the choice of method is within the competence of one skilled in the art (Robbins (ed), 1997, *Gene Therapy Protocols*, Human Press, NJ). Cells transformed with a Gene 216 gene can be used as model systems to study chromosome 20 disorders and to identify drug treatments for the treatment of such disorders.

Gene transfer systems known in the art may be useful in the practice of the gene therapy methods of the present invention. These include viral and non-viral transfer methods. A number of viruses have been used as gene transfer vectors, including polyoma, *i.e.*, SV40 (Madzak et al., 1992, *J. Gen. Virol.*, **73**:1533-1536), adenovirus (Berkner, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:39-6; Berkner et al., 1988, *Bio Techniques*, **6**:616-629; Gorziglia et al., 1992, *J. Virol.*, **66**:4407-4412; Quantin et al., 1992, *Proc. Natl. Acad. Sci. USA*, **89**:2581-2584; Rosenfeld et al., 1992, *Cell*,

68:143-155; Wilkinson et al., 1992, *Nucl. Acids Res.*, **20**:2233-2239; Stratford-Perricaudet et al., 1990, *Hum. Gene Ther.*, **1**:241-256), vaccinia virus (Mackett et al., 1992, *Biotechnology*, **24**:495-499), adeno-associated virus (Muzyczka, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:91-123; Ohi et al., 1990, *Gene*, **89**:279-282), herpes viruses including HSV and EBV (Margolskee, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:67-90; Johnson et al., 1992, *J. Virol.*, **66**:2952-2965; Fink et al., 1992, *Hum. Gene Ther.*, **3**:11-19; Breakfield et al., 1987, *Mol. Neurobiol.*, **1**:337-371; Fresse et al., 1990, *Biochem. Pharmacol.*, **40**:2189-2199), and retroviruses of avian (Brandyopadhyay et al., 1984, *Mol. Cell Biol.*, **4**:749-754; Petropoulos et al., 1992, *J. Virol.*, **66**:3391-3397), murine (Miller, 1992, *Curr. Top. Microbiol. Immunol.*, **158**:1-24; Miller et al., 1985, *Mol. Cell Biol.*, **5**:431-437; Sorge et al., 1984, *Mol. Cell Biol.*, **4**:1730-1737; Mann et al., 1985, *J. Virol.*, **54**:401-407), and human origin (Page et al., 1990, *J. Virol.*, **64**:5370-5276; Buchsachler et al., 1992, *J. Virol.*, **66**:2731-2739). Most human gene therapy protocols have been based on disabled murine retroviruses.

Non-viral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham et al., 1973, *Virology*, **52**:456-467; Pellicer et al., 1980, *Science*, **209**:1414-1422), mechanical techniques, for example microinjection (Anderson et al., 1980, *Proc. Natl. Acad. Sci. USA*, **77**:5399-5403; Gordon et al., 1980, *Proc. Natl. Acad. Sci. USA*, **77**:7380-7384; Brinster et al., 1981, *Cell*, **27**:223-231; Constantini et al., 1981, *Nature*, **294**:92-94), membrane fusion-mediated transfer via liposomes (Felgner et al., 1987, *Proc. Natl. Acad. Sci. USA*, **84**:7413-7417; Wang et al., 1989, *Biochemistry*, **28**:9508-9514; Kaneda et al., 1989, *J. Biol. Chem.*, **264**:12126-12129; Stewart et al., 1992, *Hum. Gene Ther.*, **3**:267-275; Nabel et al., 1990, *Science*, **249**:1285-1288; Lim et al., 1992, *Circulation*, **83**:2007-2011), and direct DNA uptake and receptor-mediated DNA transfer (Wolff et al., 1990, *Science*, **247**:1465-1468; Wu et al., 1991, *BioTechniques*, **11**:474-485; Zenke et al., 1990, *Proc. Natl. Acad.*

Sci. USA, **87**:3655-3659; Wu et al., 1989, *J. Biol. Chem.*, **264**:16985-16987; Wolff et al., 1991, *BioTechniques*, **11**:474-485; Wagner et al., 1991, *Proc. Natl. Acad. Sci. USA*, **88**:4255-4259; Cotten et al., 1990, *Proc. Natl. Acad. Sci. USA*, **87**:4033-4037; Curiel et al., 1991, *Proc. Natl. Acad. Sci. USA*,  
5 **88**:8850-8854; Curiel et al., 1991, *Hum. Gene Ther.*, **3**:147-154).

In one approach, plasmid DNA is complexed with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient  
10 binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

In another approach, liposome/DNA is used to mediate direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is non-specific, localized *in vivo* uptake and expression have been  
15 reported in tumor deposits, for example, following direct *in situ* administration (Nabel, 1992, *Hum. Gene Ther.*, **3**:399-410).

Suitable gene transfer vectors possess a promoter sequence, preferably a promoter that is cell-specific and placed upstream of the sequence to be expressed. The vectors may also contain, optionally, one or more expressible  
20 marker genes for expression as an indication of successful transfection and expression of the nucleic acid sequences contained in the vector. In addition, vectors can be optimized to minimize undesired immunogenicity and maximize long-term expression of the desired gene product(s) (see Nabe, 1999, *Proc. Natl. Acad. Sci. USA* **96**:324-326). Moreover, vectors can be chosen based on  
25 cell-type that is targeted for treatment. Notably, gene transfer therapies have been initiated for the treatment of various pulmonary diseases (see, e.g., M.J. Welsh, 1999, *J. Clin. Invest.* **104**(9):1165-6; D.L. Ennist, 1999, *Trends Pharmacol. Sci.* **20**:260-266; S.M. Albelda et al., 2000, *Ann. Intern. Med.* **132**:649-660; E. Alton and C. Kitson C., 2000, *Expert Opin. Investig. Drugs.*  
30 **9**(7):1523-35).

Illustrative examples of vehicles or vector constructs for transfection or

infection of the host cells include replication-defective viral vectors, DNA virus or RNA virus (retrovirus) vectors, such as adenovirus, herpes simplex virus and adeno-associated viral vectors. Adeno-associated virus vectors are single stranded and allow the efficient delivery of multiple copies of nucleic acid to the cell's nucleus. Preferred are adenovirus vectors. The vectors will normally be substantially free of any prokaryotic DNA and may comprise a number of different functional nucleic acid sequences. An example of such functional sequences may be a DNA region comprising transcriptional and translational initiation and termination regulatory sequences, including promoters (e.g., strong promoters, inducible promoters, and the like) and enhancers which are active in the host cells. Also included as part of the functional sequences is an open reading frame (polynucleotide sequence) encoding a protein of interest. Flanking sequences may also be included for site-directed integration. In some situations, the 5'-flanking sequence will allow homologous recombination, thus changing the nature of the transcriptional initiation region, so as to provide for inducible or non-inducible transcription to increase or decrease the level of transcription, as an example.

In general, the encoded and expressed Gene 216 polypeptide may be intracellular, i.e., retained in the cytoplasm, nucleus, or in an organelle, or may be secreted by the cell. For secretion, the natural signal sequence present in Gene 216 may be retained. When the polypeptide or peptide is a fragment of a Gene 216 protein, a signal sequence may be provided so that, upon secretion and processing at the processing site, the desired protein will have the natural sequence. Specific examples of coding sequences of interest for use in accordance with the present invention include the Gene polypeptide coding sequences, e.g., SEQ ID NO:4.

As previously mentioned, a marker may be present for selection of cells containing the vector construct. The marker may be an inducible or non-inducible gene and will generally allow for positive selection under induction, or without induction, respectively. Examples of marker genes include neomycin, dihydrofolate reductase, glutamine synthetase, and the like. The



vector employed will generally also include an origin of replication and other genes that are necessary for replication in the host cells, as routinely employed by those having skill in the art. As an example, the replication system comprising the origin of replication and any proteins associated with replication  
5 encoded by a particular virus may be included as part of the construct. The replication system must be selected so that the genes encoding products necessary for replication do not ultimately transform the cells. Such replication systems are represented by replication-defective adenovirus (see G. Acsadi et al., 1994, *Hum. Mol. Genet.* **3**:579-584) and by Epstein-Barr virus. Examples  
10 of replication defective vectors, particularly, retroviral vectors that are replication defective, are BAG, (see Price et al., 1987, *Proc. Natl. Acad. Sci. USA*, **84**:156; Sanes et al., 1986, *EMBO J.*, **5**:3133). It will be understood that the final gene construct may contain one or more genes of interest, for example, a gene encoding a bioactive metabolic molecule. In addition, cDNA,  
15 synthetically produced DNA or chromosomal DNA may be employed utilizing methods and protocols known and practiced by those having skill in the art.

According to one approach for gene therapy, a vector encoding a Gene 216 polypeptide is directly injected into the recipient cells (*in vivo* gene therapy). Alternatively, cells from the intended recipients are explanted,  
20 genetically modified to encode a Gene 216 polypeptide, and reimplanted into the donor (*ex vivo* gene therapy). An *ex vivo* approach provides the advantage of efficient viral gene transfer, which is superior to *in vivo* gene transfer approaches. In accordance with *ex vivo* gene therapy, the host cells are first transfected with engineered vectors containing at least one gene encoding a  
25 Gene 216 polypeptide, suspended in a physiologically acceptable carrier or excipient such as saline or phosphate buffered saline, and the like, and then administered to the host. The desired gene product is expressed by the injected cells, which thus introduce the gene product into the host. The introduced gene products can thereby be utilized to treat or ameliorate a  
30 disorder that is related to altered levels of Gene 216 (e.g., asthma).

### Animal Models

Gene 216 polynucleotides can be used to generate genetically altered non-human animals or human cell lines. Any non-human animal can be used; however typical animals are rodents, such as mice, rats, or guinea pigs.

- 5 Genetically engineered animals or cell lines can carry a gene that has been altered to contain deletions, substitutions, insertions, or modifications of the polynucleotide sequence (e.g., exon sequence). Such alterations may render the gene nonfunctional, (i.e., a null mutation) producing a "knockout" animal or cell line. In addition, genetically engineered animals can carry one or more
- 10 exogenous or non-naturally occurring genes, i.e., "transgenes", that are derived from different organisms (e.g., humans), or produced by synthetic or recombinant methods. Genetically altered animals or cell lines can be used to study Gene 216 function, regulation, and treatments for Gene 216-related diseases. In particular, knockout animals and cell lines can be used to
- 15 establish animal models and *in vitro* models for Gene 216-related illnesses, respectively. In addition, transgenic animals expressing human Gene 216 can be used in drug discovery efforts.

- A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by
- 20 deliberate genetic manipulation at a subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The term "transgenic animal" is not intended to encompass classical cross-breeding or *in vitro* fertilization, but rather is meant to encompass animals in which one or more cells are altered by, or receive, a recombinant DNA molecule. This
- 25 recombinant DNA molecule may be specifically targeted to a defined genetic locus, may be randomly integrated within a chromosome, or it may be extrachromosomally replicating DNA.

- Transgenic animals can be selected after treatment of germline cells or zygotes. For example, expression of an exogenous Gene 216 gene or a
- 30 variant can be achieved by operably linking the gene to a promoter and optionally an enhancer, and then microinjecting the construct into a zygote

(see, e.g., Hogan et al., *Manipulating the Mouse Embryo, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). Such treatments include insertion of the exogenous gene and disrupted homologous genes. Alternatively, the gene(s) of the animals may be disrupted by insertion  
5 or deletion mutation of other genetic alterations using conventional techniques (see, e.g., Capecchi, 1989, *Science*, **244**:1288; Valancuis et al., 1991, *Mol. Cell Biol.*, **11**:1402; Hasty et al., 1991, *Nature*, **350**:243; Shinkai et al., 1992, *Cell*, **68**:855; Mombaerts et al., 1992, *Cell*, **68**:869; Philpott et al., 1992, *Science*, **256**:1448; Snouwaert et al., 1992, *Science*, **257**:1083; Donehower et al., 1992,  
10 *Nature*, **356**:215).

In one aspect of the invention, Gene 216 knockout mice can be produced in accordance with well-known methods (see, e.g., M.R. Capecchi, 1989, *Science*, **244**:1288-1292; P. Li et al., 1995, *Cell* **80**:401-411; L.A. Gall-Taliadoros et al., 1995, *J. Immunol. Methods* **181**(1):1-15; C.H. Westphal et al.,  
15 1997, *Curr. Biol.* **7**(7):530-3; S.S. Cheah et al., 2000, *Methods Mol. Biol.* **136**:455-63). The disclosed murine Gene 216 genomic clone can be used to prepare a Gene 216 targeting construct that can disrupt Gene 216 in the mouse by homologous recombination at the Gene 216 chromosomal locus. The targeting construct can comprise a disrupted or deleted Gene 216  
20 sequence that inserts in place of the functioning portion of the native mouse gene. For example, the construct can contain an insertion in the Gene 216 protein-coding region.

Preferably, the targeting construct contains markers for both positive and negative selection. The positive selection marker allows the selective  
25 elimination of cells that lack the marker, while the negative selection marker allows the elimination of cells that carry the marker. In particular, the positive selectable marker can be an antibiotic resistance gene, such as the neomycin resistance gene, which can be placed within the coding sequence of Gene 216 to render it non-functional, while at the same time rendering the construct  
30 selectable. The herpes simplex virus thymidine kinase (HSV tk) gene is an example of a negative selectable marker that can be used as a second marker

to eliminate cells that carry it. Cells with the HSV tk gene are selectively killed in the presence of gangcyclovir. As an example, a positive selection marker can be positioned on a targeting construct within the region of the construct that integrates at the Gene 216 locus. The negative selection marker can be  
5 positioned on the targeting construct outside the region that integrates at the Gene 216 locus. Thus, if the entire construct is present in the cell, both positive and negative selection markers will be present. If the construct has integrated into the genome, the positive selection marker will be present, but the negative selection marker will be lost.

10 The targeting construct can be employed, for example, in embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured *in vitro* (M.J. Evans et al., 1981, *Nature* **292**:154-156; M.O. Bradley et al., 1984, *Nature* **309**:255-258; Gossler et al., 1986, *Proc. Natl. Acad. Sci. USA* **83**:9065-9069; Robertson et al., 1986, *Nature* **322**:445-448; S. A. Wood et al.,  
15 1993, *Proc. Natl. Acad. Sci. USA* **90**:4582-4584). Targeting constructs can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. Following this, the transformed ES cells can be combined with blastocysts from a non-human animal. The introduced ES cells colonize the embryo and contribute to the  
20 germ line of the resulting chimeric animal (R. Jaenisch, 1988, *Science* **240**:1468-1474). The use of gene-targeted ES cells in the generation of gene-targeted transgenic mice has been previously described (Thomas et al., 1987, *Cell* **51**:503-512) and is reviewed elsewhere (Frohman et al., 1989, *Cell* **56**:145-147; Capecchi, 1989, *Trends in Genet.* **5**:70-76; Baribault et al., 1989, *Mol. Biol. Med.* **6**:481-492; Wagner, 1990, *EMBO J.* **9**:3025-3032; Bradley et al., 1992, *Bio/Technology* **10**: 534-539).

Several methods can be used to select homologously recombined murine ES cells. One method employs PCR to screen pools of transformant cells for homologous insertion, followed by screening individual clones (Kim et al., 1988, *Nucleic Acids Res.* **16**:8887-8903; Kim et al., 1991, *Gene* **103**:227-  
30 233). Another method employs a marker gene is constructed which will only

be active if homologous insertion occurs, allowing these recombinants to be selected directly (Sedivy et al., 1989, *Proc. Natl. Acad. Sci. USA* **86**:227-231).

For example, the positive-negative selection (PNS) method can be used as described above (see, e.g., Mansour et al., 1988, *Nature* **336**:348-352; 5 Capecchi, 1989, *Science* **244**:1288-1292; Capecchi, 1989, *Trends in Genet.* **5**:70-76). In particular, the PNS method is useful for targeting genes that are expressed at low levels.

The absence of functional Gene 216 in the knockout mice can be confirmed, for example, by RNA analysis, protein expression analysis, and 10 functional studies. For RNA analysis, RNA samples are prepared from different organs of the knockout mice and the Gene 216 transcript is detected in Northern blots using oligonucleotide probes specific for the transcript. For protein expression detection, antibodies that are specific for the Gene 216 polypeptide are used, for example, in flow cytometric analysis, 15 immunohistochemical staining, and activity assays. Alternatively, functional assays are performed using preparations of different cell types collected from the knockout mice.

Several approaches can be used to produce transgenic mice. In one approach, a targeting vector is integrated into ES cell by homologous 20 recombination, an intrachromosomal recombination event is used to eliminate the selectable markers, and only the transgene is left behind (A.L. Joyner et al., 1989, *Nature* **338**(6211):153-6; P. Hasty et al., 1991, *Nature* **350**(6315):243-6; V. Valancius and O. Smithies, 1991, *Mol. Cell Biol.* **11**(3):1402-8; S. Fiering et al., 1993, *Proc. Natl. Acad. Sci. USA* **90**(18):8469-73). In an alternative 25 approach, two or more strains are created; one strain contains the gene knocked-out by homologous recombination, while one or more strains contain transgenes. The knockout strain is crossed with the transgenic strain to produce new line of animals in which the original wild-type allele has been replaced (although not at the same site) with a transgene. Notably, knockout 30 and transgenic animals can be produced by commercial facilities (e.g., The Lerner Research Institute, Cleveland, OH; B&K Universal, Inc., Fremont, CA;

DNX Transgenic Sciences, Cranbury, NJ; Incyte Genomics, Inc., St. Louis, MO).

Transgenic animals (e.g., mice) containing a nucleic acid molecule which encodes human Gene 216, may be used as *in vivo* models to study the overexpression of Gene 216. Such animals can also be used in drug evaluation and discovery efforts to find compounds effective to inhibit or modulate the activity of Gene 216, such as for example compounds for treating respiratory disorders, diseases, or conditions. One having ordinary skill in the art can use standard techniques to produce transgenic animals which produce human Gene 216 polypeptide, and use the animals in drug evaluation and discovery projects (see, e.g., U.S. Patent No. 4,873,191 to Wagner; U.S. Patent No. 4,736,866 to Leder).

In another embodiment of the present invention, the transgenic animal can comprise a recombinant expression vector in which the nucleotide sequence that encodes human Gene 216 is operably linked to a tissue specific promoter whereby the coding sequence is only expressed in that specific tissue. For example, the tissue specific promoter can be a mammary cell specific promoter and the recombinant protein so expressed is recovered from the animal's milk.

In yet another embodiment of the present invention, a Gene 216 "knockout" can be produced by administering to the animal antibodies (e.g., neutralizing antibodies) that specifically recognize an endogenous Gene 216 polypeptide. The antibodies can act to disrupt function of the endogenous Gene 216 polypeptide, and thereby produce a null phenotype. In one specific example, an orthologous mouse Gene 216 polypeptide (e.g., SEQ ID NO:366) or peptide can be used to generate antibodies. These antibodies can be given to a mouse to knockout the function of the mouse Gene 216 ortholog.

In addition, non-mammalian organisms may be used to study Gene 216 and Gene 216-related diseases. For example, model organisms such as *C. elegans*, *D. melanogaster*, and *S. cerevisiae* may be used. Gene 216 homologues can be identified in these model organisms, and mutated or

deleted to produce a Gene 216-deficient strain. Human Gene 216 can then be tested for the ability to "complement" the Gene 216-deficient strain. Gene 216-deficient strains can also be used for drug screening. The study of Gene 216 homologs can facilitate the understanding of human Gene 216 biological function, and assist in the identification of binding proteins (e.g., agonists and antagonists).

### **Gene Identification**

To identify genes in the region on 20p13-p12, a set of bacterial artificial chromosome(BAC) clones containing this chromosomal region was identified in accordance with the methods described herein. The BAC clones served as a template for genomic DNA sequencing and served as reagents for identifying coding sequences by direct cDNA selection. Genomic sequencing and direct cDNA selection methods were used to characterize DNA from 20p13-p12.

When one or more genes have been genetically localized to a specific chromosomal region, the gene(s) can be characterized at the molecular level by a series of steps that include: 1) cloning the entire region of DNA in a set of overlapping clones (physical mapping); 2) characterizing the gene(s) encoded by these clones by a combination of direct cDNA selection, exon trapping and DNA sequencing (gene identification); and 3) identifying mutations (i.e., SNPs) in the gene(s) by comparative DNA sequencing of affected and unaffected members of the kindred and/or in unrelated affected individuals and unrelated unaffected controls (mutation analysis).

Physical mapping is accomplished by screening libraries of human DNA cloned in vectors that are propagated in a host such as *E. coli*, using hybridization or PCR assays from unique molecular landmarks in the chromosomal region of interest. In accordance with the present invention, a physical map of the disorder region was generated by screening a library of human DNA cloned in BACs with a set overgo markers that had been previously mapped to chromosome 20p13-p12 by the efforts of the Human Genome Project. Overgos are unique molecular landmarks in the human genome that can be assayed by hybridization. The location of thousands of

overgos on the twenty-two autosomes and two sex chromosomes has been determined through the efforts of the Human Genome Project. For a positional cloning effort, the physical map is tied to the genetic map because the markers used for genetic mapping can also be used as overgos for physical mapping.

- 5 By screening a BAC library with a combination of overgos derived from genetic markers, genes, and random DNA fragments, a physical map comprised of overlapping clones representing all of the DNA in a chromosomal region of interest can be assembled.

BACs are cloning vectors for large (80 kilobase to 200 kilobase) segments of human or other DNA that are propagated in *E. coli*. To construct a physical map using BACs, a library of BAC clones is screened so that individual clones harboring the DNA sequence corresponding to a given overgo or set of overgos are identified. Throughout most of the human genome, the overgo markers are spaced approximately 20 to 50 kilobases apart, so that an individual BAC clone typically contains at least two overgo markers. In addition, the BAC libraries that were screened contain enough cloned DNA to cover the human genome twelve times over. An individual overgo typically identifies more than one BAC clone. By screening a twelve-fold coverage BAC library with a series of overgo markers spaced approximately 50 kilobases apart, a physical map consisting of a series of overlapping contiguous BAC clones, i.e., BAC "contigs," can be assembled for any region of the human genome. This map is closely tied to the genetic map because many of the overgo markers used to prepare the physical map are also genetic markers.

When constructing a physical map, it often happens that there are gaps in the overgo map of the genome that result in the inability to identify BAC clones that are overlapping in a given location. Typically, the physical map is first constructed from a set of overgos identified through the publicly available literature and World Wide Web resources. The initial map consists of several separate BAC contigs that are separated by gaps of unknown molecular distance. To identify BAC clones that fill these gaps, it is necessary to develop new overgo markers from the ends of the clones on either side of the gap.



This is done by sequencing the terminal 200 to 300 base pairs of the BACs flanking the gap, and developing a PCR or hybridization based assay. If the terminal sequences are demonstrated to be unique within the human genome, then the new overgo can be used to screen the BAC library to identify  
5 additional BACs that contain the DNA from the gap in the physical map. To assemble a BAC contig that covers a region the size of the disorder region (6,000,000 or more base pairs), it is necessary to develop new overgo markers from the ends of a number of clones.

After building a BAC contig, this set of overlapping clones serves as a  
10 template for identifying the genes encoded in the chromosomal region. Gene identification can be accomplished by many methods. Three methods are commonly used: 1) a set of BACs selected from the BAC contig to represent the entire chromosomal region are sequenced, and computational methods are used to identify all of the genes; 2) the BACs from the BAC contig are used as  
15 a reagent to clone cDNAs corresponding to the genes encoded in the region by a method termed direct cDNA selection; or 3) the BACs from the BAC contig are used to identify coding sequences by selecting for specific DNA sequence motifs in a procedure called exon trapping. Gene 216 was identified by methods (1) and (2) in accordance with the techniques disclosed herein.

20 To sequence the entire BAC contig representing the disorder region, a set of BACs can be chosen for subcloning into plasmid vectors and subsequent DNA sequencing of these subclones. Since the DNA cloned in the BACs represents genomic DNA, this sequencing is referred to as genomic sequencing to distinguish it from cDNA sequencing. To initiate the genomic  
25 sequencing for a chromosomal region of interest, several non-overlapping BAC clones are chosen. DNA for each BAC clone is prepared, and the clones are sheared into random small fragments that are subsequently cloned into standard plasmid vectors such as pUC18. The plasmid clones are then grown to propagate the smaller fragments, and these are the templates for  
30 sequencing. To ensure adequate coverage and sequence quality for the BAC DNA sequence, sufficient plasmid clones are sequenced to yield three-fold

coverage of the BAC clone. For example, if the BAC is 100 kilobases long, then phagemids are sequenced to yield 300 kilobases of sequence. Since the BAC DNA is randomly sheared prior to cloning in the phagemid vector, the 300 kilobases of raw DNA sequence can be assembled by computational methods  
5 into overlapping DNA sequences termed sequence contigs. For the purposes of initial gene identification by computational methods, three-fold coverage of each BAC is sufficient to yield twenty to forty sequence contigs of 1000 base pairs to 20,000 base pairs.

In accordance with the present invention, the "seed" BACs from the  
10 BAC contig in the disorder region were sequenced. The sequence of the "seed" BACs was then used to identify minimally overlapping BACs from the contig, and these were subsequently sequenced. In this manner, the entire candidate region can be sequenced, with several small sequence gaps left in each BAC. This sequence serves as the template for computational gene  
15 identification. In one approach, genes can be identified by comparing the sequence of BAC contig to publicly available databases of cDNA and genomic sequences, e.g. UniGene, dbEST, EMBL nucleotide database, GenBank, and the DNA Database of Japan (DDBJ). The BAC DNA sequence can also be translated into protein sequence, and the protein sequence can be used to  
20 search publicly available protein databases, e.g., GenPept, EMBL protein database, Protein Information Resource (PIR), Protein Data Bank (PDB), and SWISS-PROT. These comparisons are typically done using the BLAST family of computer algorithms and programs (Altschul et al., 1990, *J. Mol. Biol.*, 215:403-410; Altschul et al, 1997, *Nucl. Acids Res.*, **25**:3389-3402).

25 For nucleotide queries, BLASTN, BLASTX, and TBLASTX can be used. BLASTN compares a nucleotide query sequence with a nucleotide sequence database; BLASTX compares a nucleotide query sequence translated in all reading frames against a protein sequence database; TBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame  
30 translations of a nucleotide sequence database. For protein queries, BLASTP and TBLASTN can be used. BLASTP compares a protein query sequence

with a protein sequence database; TBLASTN compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames.

Additionally, computer algorithms such as MZEF (Zhang, 1997, *Proc. Natl. Acad. Sci. USA* **94**:565-568), GAIL (Uberbacher et al., 1996, *Methods Enzymol.*, **266**:259-281), and Genscan (Burge and Karlin, 1997, *J. Mol. Biol.*, **268**:78-94) can be used to predict the location of exons in the sequence based on the presence of specific DNA sequence motifs that are common to all exons, as well as the presence of codon usage typical of human protein encoding sequences.

In addition to identifying genes by computational methods, genes can be identified by direct cDNA selection (DeI Mastro and Lovett, 1996, *Methods in Molecular Biology*, Humana Press Inc., NJ). In direct cDNA selection, cDNA pools from tissues of interest are prepared, and BACs from the candidate region are used in a liquid hybridization assay to capture the cDNAs which base pair to coding regions in the BAC. In the methods described herein, the cDNA pools were created from several different tissues by random priming and oligo dT priming the first strand cDNA from poly A<sup>+</sup> RNA, synthesizing the second-strand cDNA by standard methods, and adding linkers to the ends of the cDNA fragments. In this approach, the linkers are used to amplify the cDNA pools of BAC clones from the disorder region identified by screening a BAC library. The amplified products are then used as a template for initiating DNA synthesis to create a biotin labeled copy of BAC DNA. Following this, the biotin labeled copy of the BAC DNA is denatured and incubated with an excess of the PCR amplified, linkered cDNA pools which have also been denatured. The BAC DNA and cDNA are allowed to anneal in solution, and heteroduplexes between the BAC and the cDNA are isolated using streptavidin coated magnetic beads. The cDNAs that are captured by the BAC are then amplified using primers complimentary to the linker sequences, and the hybridization/selection process is repeated for a second round. After two rounds of direct cDNA selection, the cDNA fragments are cloned, and a library

of these direct selected fragments is created.

The cDNA clones isolated by direct selection are analyzed by two methods. Where the genomic target DNA sequence is obtained from a pool of BACs from the disorder region, the cDNAs are mapped to BAC genomic clones to verify their chromosomal location. This is accomplished by arraying the cDNAs in microtiter dishes, and replicating their DNA in high-density grids. Individual genomic clones known to map to the region are then hybridized to the grid to identify direct selected cDNAs mapping to that region. cDNA clones that are confirmed to correspond to individual BACs are sequenced. To determine whether the cDNA clones isolated by direct selection share sequence identity or similarity to previously identified genes, the DNA and protein coding sequences are compared to publicly available databases using the BLAST family of programs described above.

The combination of genomic DNA sequence and cDNA sequence provided by BAC sequencing and by direct cDNA selection yields an initial list of putative genes in the region. In the present invention, the genes in the region were candidates for the asthma locus. To further characterize each gene, Northern blots were performed to determine the size of the transcript corresponding to each gene, and to determine which putative exons were transcribed together to make an individual gene. For Northern blot analysis of each gene, probes are prepared from direct selected cDNA clones or by PCR amplifying specific fragments from genomic DNA, cDNA or from the BAC encoding the putative gene of interest. The Northern blot analysis is used to determine the size of the transcript and the tissues in which it is expressed. For transcripts that are not highly expressed, it is sometimes necessary to perform a reverse transcription PCR assay using RNA from the tissues of interest as a template for the reaction.

Gene identification by computational methods and by direct cDNA selection provides unique information about the genes in a region of a chromosome. Once genes are identified, it is possible to examine subjects for sequence variants. Variant sequences can be inherited as allelic differences

or can arise from spontaneous mutations.

Inherited alleles can be analyzed for linkage to a disease susceptibility locus. Linkage analysis is possible because of the nature of inheritance of chromosomes from parents to offspring. During meiosis, the two parental  
5 homologs pair to guide their proper separation to daughter cells. While they are paired, the two homologs exchange pieces of the chromosomes, in an event called "crossing over" or "recombination." The resulting chromosomes contain parts that originate from both parental homologs. The closer together two sequences are on the chromosome, the less likely that a recombination  
10 event will occur between them, and the more closely linked they are.

In the present invention, data obtained from the different families were combined and analyzed together by a computer using statistical methods described herein. The results were then used as evidence for linkage between the genetic markers used and an asthma susceptibility locus.

15 In general, a recombination frequency of 1% is equivalent to approximately 1 map unit, a relationship that holds up to frequencies of about 20% or 20 cM. One centimorgan (cM) is roughly equivalent to 1,000 kb of DNA. The entire human genome is 3,300 cM long. In order to find an unknown disease gene within 5-10 cM of a marker locus, the whole human  
20 genome can be searched with roughly 330 informative marker loci spaced at approximately 10 cM intervals (Botstein et al., 1980, *Am. J. Hum. Genet.*, **32**:314-331).

The reliability of linkage results is established by using a number of statistical methods. The methods most commonly used for the detection by  
25 linkage analysis of oligogenes involved in the etiology of a complex trait are non-parametric or model-free methods which have been implemented into the computer programs MAPMAKER/SIBS (L. Kruglyak and E.S. Lander, 1995, *Am. J. Hum. Genet.* **57**:439-454) and GENEHUNTER (L. Kruglyak et al., 1996, *Am. J. Hum. Genet.* **58**:1347-1363). Typically, linkage analysis is performed  
30 by typing members of families with multiple affected individuals at a given marker locus and evaluating if the affected members (excluding parent-

offspring pairs) share alleles at the marker locus that are identical by descent (IBD) more often than expected by chance alone.

As a result of the rapid advances in mapping the human genome over the last few years, and concomitant improvements in computer methodology, it has become feasible to carry out linkage analyses using multi-point data. Multi-point analysis provides a simultaneous analysis of linkage between the trait and several linked genetic markers, when the recombination distance among the markers is known. A LOD score statistic is computed at multiple locations along a chromosome to measure the evidence that a susceptibility locus is located nearby. A LOD score is the logarithm base 10 of the ratio of the likelihood that a susceptibility locus exists at a given location to the likelihood that no susceptibility locus is located there. By convention, when testing a single marker, a total LOD score greater than +3.0 (that is, odds of linkage being 1,000 times greater than odds of no linkage) is considered to be significant evidence for linkage.

Multi-point analysis is advantageous for two reasons. First, the informativeness of the pedigrees is usually increased. Each pedigree has a certain amount of potential information, dependent on the number of parents heterozygous for the marker loci and the number of affected individuals in the family. However, few markers are sufficiently polymorphic as to be informative in all those individuals. If multiple markers are considered simultaneously, then the probability of an individual being heterozygous for at least one of the markers is greatly increased. Second, an indication of the position of the disease gene among the markers may be determined. This allows identification of flanking markers, and thus eventually allows identification of a small region in which the disease gene resides.

## EXAMPLES

The examples as set forth herein are meant to exemplify the various aspects of the present invention and are not intended to limit the invention in any way.

### 5 **EXAMPLE 1: Family Collection**

Asthma is a complex disorder that is influenced by a variety of factors, including both genetic and environmental effects. Complex disorders are typically caused by multiple interacting genes, some contributing to disease development and some conferring a protective effect. The success of linkage  
10 analyses in identifying chromosomes with significant LOD scores is achieved in part as a result of an experimental design tailored to the detection of susceptibility genes in complex diseases, even in the presence of epistasis and genetic heterogeneity. Also important are rigorous efforts in ascertaining asthmatic families that meet strict guidelines, and collecting accurate clinical  
15 information.

Given the complex nature of the asthma phenotype, non-parametric affected sib pair analyses were used to analyze the genetic data. This approach does not require parameter specifications such as mode of inheritance, disease allele frequency, penetrance of the disorder, or phenocopy  
20 rates. Instead, it determines whether the inheritance pattern of a chromosomal region is consistent with random segregation. If it is not, affected sibs inherit identical copies of alleles more often than expected by chance. Because no models for inheritance are assumed, allele-sharing methods tend to be more robust than parametric methods when analyzing complex disorders. They do,  
25 however, require larger sample sizes to reach statistically significant results.

At the outset of the program, the goal was to collect 400 affected sib-pair families for the linkage analyses. Based on a genome scan with markers spaced ~10 cM apart, this number of families was predicted to provide > 95% power to detect an asthma susceptibility gene that caused an increased risk  
30 to first-degree relatives of 3-fold or greater. The assumed relative risk of 3-fold was consistent with epidemiological studies in the literature that suggest an

increased risk ranging from 3- to 7-fold. The relative risk was based on gender, different classifications of the asthma phenotype (i.e. bronchial hyper-responsiveness versus physician's diagnosis) and, in the case of offspring, whether one or both parents were asthmatic.

5           The family collection efforts exceeded the initial goal of 400, obtaining a total of 444 affected sibling pair (ASP) families, with 342 families from the UK and 102 families from the US. The ASP families in the US collection were Caucasian with a minimum of two affected siblings that were identified through both private practice and community physicians as well as through advertising.

10          A total of 102 families were collected in Kansas, Nebraska, and Southern California. In the UK collection, Caucasian families with a minimum of two affected siblings were identified through physicians' registers in a region surrounding Southampton and including the Isle of Wight. In both the US and UK collections, additional affected and unaffected sibs were collected

15          whenever possible. An additional 39 families from the United Kingdom were utilized from an earlier collection effort with different ascertainment criteria. These families were recruited either: 1) without reference to asthma and atopy; or 2) by having at least one family member or at least two family members affected with asthma. The randomly ascertained samples were

20          identified from general practitioner registers in the Southampton area. For families with affected members, the probands were recruited from hospital based clinics in Southampton. Seven pedigrees extended beyond a single nuclear family.

            Families were included in the study if they met all of the following

25          criteria: 1) the biological mother and biological father were Caucasian and agreed to participate in the study; 2) at least two biological siblings were alive, each with a current physician diagnosis of asthma, and were 5 to 21 years of age; and 3) the two siblings were currently taking asthma medications on a regular basis. This included regular, intermittent use of inhaled or oral

30          bronchodilators and regular use of cromolyn, theophylline, or steroids.

            Families were excluded from the study if they met any one of the



following criteria: 1) both parents were affected (i.e., with a current diagnosis of asthma, having asthma symptoms, or on asthma medications at the time of the study); 2) any of the siblings to be included in the study was less than 5 years of age; 3) any asthmatic family member to be included in the study was taking beta-blockers at the time of the study, 4) any family member to be included in the study had congenital or acquired pulmonary disease at birth (e.g. cystic fibrosis), a history of serious cardiac disease (myocardial infarction) or any history of serious pulmonary disease (e.g. emphysema); or 5) any family member to be included in the study was pregnant.

10       An extensive clinical instrument was designed and data from all participating family members were collected. The case report form (CRF) included questions on demographics, medical history including medications, a health survey on the incidence and frequency of asthma, wheeze, eczema, hay fever, nasal problems, smoking, and questions on home environment.

15       Data from a video questionnaire designed to show various examples of wheeze and asthmatic attacks were also included in the CRF. Clinical data, including skin prick tests to 8 common allergens, total and specific IgE levels, and bronchial hyper-responsiveness following a methacholine challenge, were also collected from all participating family members. All data were entered into a

20       SAS dataset by IMTCI, a CRO; either by double data entry or scanning followed by on-screen visual validation. An extensive automated review of the data was performed on a routine basis and a full audit at the conclusion of the data entry was completed to verify the accuracy of the dataset.

#### **EXAMPLE 2: Genome Scan**

25       In order to identify chromosomal regions linked to asthma, the inheritance pattern of alleles from genetic markers spanning the genome was assessed on the collected family resources. As described above, combining these results with the segregation of the asthma phenotype in these families allows the identification of genetic markers that are tightly linked to asthma. In

30       turn, this provides an indication of the location of genes predisposing affected individuals to asthma. The genotyping strategy was twofold: 1) to conduct a

genome wide scan using markers spaced at approximately 10 cM intervals; and 2) to target ten chromosomal regions for high density genetic mapping.

The initial candidate regions for high-density mapping were chosen based on suggestions of linkage to these regions by other investigators.

5            Genotypes of PCR amplified simple sequence microsatellite genetic linkage markers were determined using ABI model 377 Automated Sequencers (PE Applied Biosystems). Microsatellite markers were obtained from Research Genetics Inc. (Huntsville, AL) in the fluorescent dye-conjugated form (see Dubovsky et al., 1995, *Hum. Mol. Genet.* 4(3):449-452). The markers  
10           comprised a variation of a human linkage mapping panel as released from the Cooperative Human Linkage Center (CHLC), also known as the Weber lab screening set version 8. The variation of the Weber 8 screening set consisted of 529 markers with an average spacing of 6.9 cM (autosomes only) and 7.0 cM (all chromosomes). Eighty-nine percent of the markers consisted of either  
15           tri- or tetra-nucleotide microsatellites. There were no gaps present in chromosomal coverage greater than 17.5 cM.

Study subject genomic DNA (5 µl; 4.5 ng/µl) was amplified in a 10 µl PCR reaction using AmpliTaq Gold DNA polymerase (0.225 U); 1 X PCR buffer (80 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 30 mM Tris-HCl (pH 8.8); 0.5% Tween-20); 200 µM each  
20           dATP, dCTP, dGTP and dTTP; 1.5-3.5 µM MgCl<sub>2</sub>; and 250 µM forward and reverse PCR primers. PCR reactions were set up in 192 well plates (Costar) using a Tecan Genesis 150 robotic workstation equipped with a refrigerated deck. PCR reactions were overlaid with 20 µl mineral oil, and thermocycled on an MJ Research Tetrad DNA Engine equipped with four 192 well heads using  
25           the following conditions: 92°C for 3 min; 6 cycles of 92°C for 30 sec, 56°C for 1 min, 72°C for 45 sec; followed by 20 cycles of 92°C for 30 sec, 55°C for 1 min, 72°C for 45 sec; and a 6 min incubation at 72°C.

PCR products of 8-12 microsatellite markers were subsequently pooled into two 96-well microtitre plates (2.0 µl PCR product from TET and FAM  
30           labeled markers, 3.0 µl HEX labeled markers) using a Tecan Genesis 200 robotic workstation and brought to a final volume of 25 µl with H<sub>2</sub>O. Following

this, 1.9  $\mu$ l of pooled PCR product was transferred to a loading plate and combined with 3.0  $\mu$ l loading buffer (2.5  $\mu$ l formamide/blue dextran (9.0 mg/ml), 0.5  $\mu$ l GS-500 TAMRA labeled size standard, ABI). Samples were denatured in the loading plate for 4 min at 95°C, placed on ice for 2 min, and  
5 electrophoresed on a 5% denaturing polyacrylamide gel (FMC on the ABI 377XL). Samples (0.8  $\mu$ l) were loaded onto the gel using an 8 channel Hamilton Syringe pipettor.

Each gel consisted of 62 study subjects and 2 control subjects (CEPH parents ID #1331-01 and 1331-02, Coriell Cell Repository, Camden, NJ).  
10 Genotyping gels were scored in duplicate by investigators blind to patient identity and affection status using GENOTYPER analysis software V 1.1.12 (ABI; PE Applied Biosystems). Nuclear families were loaded onto the gel with the parents flanking the siblings to facilitate error detection. The final tables obtained from the GENOTYPER output for each gel analysed were imported  
15 into a SYBASE Database.

Allele calling (binning) was performed using the SYBASE version of the ABAS software (Ghosh et al., 1997, *Genome Research* 7:165-178). Offsize bins were checked manually and incorrect calls were corrected or blanked. The binned alleles were then imported into the program MENDEL (Lange et al.,  
20 1988, *Genetic Epidemiology*, 5:471) for inheritance checking using the USERM13 subroutine (Boehnke et al., 1991, *Am. J. Hum. Genet.* 48:22-25). Non-inheritance was investigated by examining the genotyping traces and, once all discrepancies were resolved, the subroutine USERM13 was used to estimate allele frequencies.

### 25 **EXAMPLE 3: Linkage Analysis**

Chromosomal regions harboring asthma susceptibility genes by linkage analysis of genotyping data and three separate phenotypes, asthma, bronchial hyper-responsiveness, and atopic status were identified as follows.

1. Asthma Phenotype: For the initial linkage analysis, the  
30 phenotype and asthma affection status were defined by a patient who answered the following questions in the affirmative: i) have you ever had

asthma; ii) do you have a current physician's diagnosis of asthma; and iii) are you currently taking asthma medications? Medications included inhaled or oral bronchodilators, cromolyn, theophylline, or steroids. Multipoint linkage analyses of allele sharing in affected individuals were performed using the

5 MAPMAKER/SIBS analysis program (L. Kruglyak and E.S. Lander, 1995, *Am. J. Hum. Genet.* **57**:439-454). The map location and distances between markers were obtained from the genetic maps published by the Marshfield medical research foundation (<http://www.marshmed.org/genetics/>). Ambiguous ordering of markers in the Marshfield map was resolved using the program

10 MULTIMAP (T.C. Matise et al., 1994, *Nature Genet.* **6**:384-390).<sup>6</sup>

Using the discrete phenotype of asthma (yes/no), a candidate region was identified on chromosome 20 with a LOD score of 2.94, based on 462 nuclear families. Figure 1 displays the multipoint LOD score against the map location of the markers along chromosome 20. A Maximum LOD Score (MLS)

15 of 2.94 was obtained at location 7.9 cM, 0.3 cM proximal to marker D20S906. A second MLS of 2.94 was obtained at marker D20S482 at location 12.1 cM. An excess sharing by descent (Identity By Descent (IBD) = 2) of 0.31 was observed at both maximum LOD scores. Table 2 lists the single and multipoint LOD scores at each marker. Analyses were done using a conservative

20 approach by weighting multiple sibling pairs within a sibship. When affected sib pairs were utilized in the linkage analyses without weighting the LOD score on chromosome 20 maximized at D20S482 with a value of 3.19. Thus, these data provided strong evidence for the presence of an asthma susceptibility gene in this region of chromosome 20.

25

**TABLE 2**

Marker	Distance	Single-point	Multipoint
D20S502	0.5	0.7	2.4
D20S103	2.1	2.4	2.3
D20S117	2.8	1.2	2.0
GTC4ATG	6.3	2.4	2.5
GTC3CA	6.6	1.3	2.7
D20S906	7.6	2.9	2.9
D20S842	9.0	1.3	2.5
D20S181	9.5	1.8	2.6
D20S193	9.5	2.5	2.5

D20S889	11.2	1.6	2.6
D20S482	12.1	1.9	2.9
D20S849	14.0	0.8	2.0
D20S835	15.1	0.5	1.8
D20S448	18.8	1.4	1.4
D20S602	21.2	1.1	1.1
D20S851	24.7	1.0	0.8
D20S604	32.9	0.0	0.1
D20S470	39.3	0.0	0.1
D20S477	47.5	0.0	0.0
D20S478	54.1	0.0	0.0
D20S481	62.3	0.0	0.0
D20S480	79.9	0.0	0.0
D20S171	95.7	0.4	0.1

2. Phenotypic Subgroups: Nuclear families were ascertained by the presence of at least two affected siblings with a current physician's diagnosis of asthma, as well as the use of asthma medication. In the initial analysis (see above), the evidence was examined for linkage based on that dichotomous phenotype (asthma – yes/no). To further characterize the linkage signals, additional quantitative traits were measured in the clinical protocol. Since quantitative trait loci (QTL) analysis tools with correction for ascertainment was not available, the following approach was taken to refine the linkage and association analyses:

i. Phenotypic subgroups that could be indicative of an underlying genotypic heterogeneity were identified. Asthma subgroups were defined according to 1) bronchial hyper-responsiveness (BHR) to methacholine challenge; or 2) to atopic status using quantitative measures like total serum IgE and specific IgE to common allergens.

ii. Non-parametric linkage analyses were performed on subgroups to test for the presence of a more homogeneous sub-sample. If genetic heterogeneity was present in the sample, the amount of allele sharing among phenotypically similar siblings was expected to increase in the appropriate subgroup in comparison to the full sample. A narrower region of significant increased allele sharing was also expected to result unless the overall LOD score decreased as a consequence of having a smaller sample size and of using an approximate partitioning of the data.

iii. Alternatively, allele sharing probabilities were

parameterized as a function of the quantitative trait value of each child in a given sib pair, as advocated by N. Morton and implemented in his program BETA (N. Morton, 1996, *Proc. Natl. Acad. Sci. USA* **93**:3471-3476). This approach alleviated the need to dichotomize a quantitative trait. However, the  
5 program did not correct for the use of non-independent sib pairs in sibship of size 3 or larger. As such it did not provide an accurate measure of the significance of a linkage finding, but was used to corroborate the localization of the linkage signal.

3. Results for BHR and IgE: PC<sub>20</sub>, the concentration of  
10 methacholine resulting in a 20% drop in FEV<sub>1</sub> (forced expiratory volume), was polychotomized in four groups and analyses were performed on the subsets of asthmatic children with mild to severe BHR (PC<sub>20</sub> ≤ 4 mg/ml) or PC<sub>20</sub>(4), as well as on the broader subset with borderline to severe BHR (PC<sub>20</sub> ≤ 16 mg/ml) or PC<sub>20</sub>(16). As shown in the LOD plot in Figure 2, the MLS for the subset of  
15 127 nuclear families with at least two PC<sub>20</sub>(4) affected sibs was 2.97 at 11.8 cM, 0.3 cM from D20S482, with an excess sharing by descent of 0.37. As shown in Figure 3, for the 218 nuclear families with at least two PC<sub>20</sub>(16), the MLS was 3.93 at D20S482 with an excess sharing of 0.36. Both PC<sub>20</sub>(4) and PC<sub>20</sub>(16) strongly implicated the region of chromosome 20 under the second  
20 peak around marker D20S482. When considering the more extreme phenotype, PC<sub>20</sub>(4), a higher proportion of families was linked to the region. However, the increase in LOD score for the PC<sub>20</sub>(16) phenotype indicated that families concordant for the milder BHR phenotype also contributed to the linkage signal and would provide a larger pool of linked families.

25 Total IgE was dichotomized using an age specific cutoff for elevated levels (one standard deviation above the mean). Similarly, a dichotomous variable was created using specific IgE to common allergens. An individual was assigned a high specific IgE value if his/her level was positive (grass or tree) or elevated (> 0.35 KU/L for cat, dog, mite A, mite B, alternaria, or  
30 ragweed) for at least one such measure. In linkage analyses, the subset of asthmatic children with high total IgE (274 families) was given a maximum LOD

score of 2.3 at 11.6 cM (Figure 4), while the subset with high specific IgE (288 families) was given a LOD score of 1.87 at 12.1 cM (Figure 5). Similar to the BHR results, analyses based on IgE implicated the region under the second peak around marker D20S482. The substantially lower LOD scores using the  
5 subset of affected sibs concordant for atopy indicated the presence of groups with fewer linked families. Thus, atopy in asthmatic individuals was not the primary phenotype associated with the linkage signal on chromosome 20.

The BETA program (Morton, 1996) was used on two scales for  $PC_{20}$ . Individuals that did not drop 20% by the last dose administered (16 mg/ml) were assigned an arbitrary value of 32 mg/ml. First, a (0,1)-severity scale was  
10 constructed by applying a linear transformation to  $PC_{20}$  where 0 mg/ml received a score of 1 and 32 mg/ml received a score of 0. For this scale, individuals that did not drop 20% in their  $FEV_1$  did not contribute to the LOD score. A maximum LOD score of 3.43 was achieved at 12.1 cM with marker  
15 D20S482. Second, a linear transformation of  $PC_{20}$  was used where 0 mg/ml received a score of 1 and 32 mg/ml a score of -1. In other words, in addition to the high concordant pairs, discordant pairs and concordant pairs that did not drop would also contribute to the LOD score. In contrast, individuals with  $PC_{20}$  close to 16 mg/ml would have little impact on the LOD score. A maximum LOD  
20 score of 2.08 was again achieved at 12.1 cM.

Accordingly, a consistent pattern of evidence by linkage analysis pointed to the existence of an asthma susceptibility locus in the vicinity of marker D20S482. This was supported by the initial analysis of the asthma (yes/no) phenotype and by analyses of BHR in asthmatic individuals. Localization in the  
25 region of marker D20S482 was obtained using both BHR and IgE phenotypes.

#### **EXAMPLE 4: Physical Mapping**

The linkage results for chromosome 20 described above were used to delineate a candidate region for a disorder-associated gene located on chromosome 20. Gene discovery efforts were thus initiated in a 25 cM interval  
30 from the 20p telomere (marker D20S502) to marker D20S851, representing a >98% confidence interval. All genes known to map to this interval were

considered as candidates. Intensive physical mapping (BAC contig construction) focused on a 90% confidence interval between markers D20S103 and D20S916, a 15 cM interval. The discovery of novel genes using direct cDNA selection focused on a 95% confidence interval between markers  
5 D20S502 (20p telomere) and D20S916, a 17 cM region.

The following section describes details of the efforts to generate cloned coverage of the disorder gene region on chromosome 20, i.e., construction of a BAC contig spanning the region. There were two primary reasons for using this approach: 1) to provide genomic clones for DNA sequencing (analysis of  
10 this sequence would provide information about the gene content of the region); and 2) to provide reagents for direct cDNA selection (this would provide additional information about novel genes mapping to the interval). The physical map consisted of an ordered set of molecular landmarks, and a set of bacterial artificial chromosome clones (BACs; U.-J. Kim et al., 1996,  
15 *Genomics* **34**:213-218; H. Shizuya et al., 1992, *Proc. Natl. Acad. Sci. USA* **89**:8794-8797) that contained the disorder gene region from human chromosome 20p13-p12.

Figure 6 depicts the BAC/STS content contig map of human chromosome 20p13-p12. Markers used to screen the RPCI-11 BAC library (P.  
20 deJong, Roswell Park Cancer Institute (RPCI)) are shown in the top row. Markers that were present in the Genome Database (GDB, <http://gdbwww.gdb.org/>) are represented by GDB nomenclature. The BAC clones are shown below the markers as horizontal lines. BAC RPCI-11\_1098L22 is labeled and the location of Gene 216, described herein, is  
25 indicated at the top of the figure.

1. Map Integration. Various publicly available mapping resources were utilized to identify existing STS (sequence tagged site) markers (Olson et al., 1989, *Science*, **245**:1434-1435) in the 20p13-p12 region. Resources included the GDB (<http://gdbwww.gdb.org/>), Genethon ([http://www.genethon.fr/genethon\\_en.html](http://www.genethon.fr/genethon_en.html)), Marshfield Center for Medical Genetics  
30 (<http://www.marshmed.org/genetics/>), the Whitehead Institute Genome Center



(<http://www-genome.wi.mit.edu/>), GeneMap98, dbSTS and dbEST (NCBI, <http://www.ncbi.nlm.nih.gov/>), the Sanger Centre (<http://www.sanger.ac.uk/>), and the Stanford Human Genome Center (<http://www-shgc.stanford.edu/>). Maps were integrated manually to identify markers mapping to the disorder region. A list of the markers is provided in Table 3.

2. Marker Development: Sequences for existing STSs were obtained from the GDB, RHDB (<http://www.ebi.ac.uk/RHdb/>), or NCBI, and were used to pick primer pairs (overgos; see Table 3) for BAC library screening. Novel markers were developed either from publicly available genomic sequences, proprietary cDNA sequences, or from sequences derived from BAC insert ends (described below). Primers were chosen using a script that automatically performs vector and repetitive sequence masking using CROSSMATCH (P. Green, University of Washington). Subsequent primer selection was performed using a customized Filemaker Pro database (<http://www.filemaker.com>). Primers for use in PCR-based clone confirmation or radiation hybrid mapping (described below) were chosen using the program Primer3 (Steve Rozen, Helen J. Skaletsky, 1996, 1997, [http://www-genome.wi.mit.edu/genome\\_software/other/primer3.html](http://www-genome.wi.mit.edu/genome_software/other/primer3.html)).

**Table 3**

Overgo	Locus	DNA Type	Gene	Forward Primer	SEQ ID NO	Reverse Primer	SEQ ID NO
stSG24277		Genomic		aactctgaaatgagaagcgtg	34	aaccaccacggattcacgcttc	45
stSG408		EST		aatatcatgcaccatgacccac	35	ataaccagatggctgtgggtca	46
A005005		EST	Attractin (ATTN)	tgaggtaagtattgtaactat	36	atccccgaatgaaatagitta	47
B849D17AL		BACend		ggagcttatcctggattatcta	37	gttgagagcccacttagataat	48
SN2		EST	Sialoadhesin (SN)	agagccacacatccatgtcctg	38	gcattgggggaagccaggacat	49
AFMb026xh5	D20S867	MSAT		aagccactctgtgaattgcat	39	gccactaggaggcaatggcaat	50
SN1		EST	Sialoadhesin (SN)	gagtagtcgtatgaccagatgg	40	cgaaggcatcacggocatctgg	51
stsH22126		EST		gtctggcaatggagcatgaaaa	41	tccagggtcattcatttcacg	52
W14876	D20S752	Genomic		attagagcacatgaaggaaagg	42	tgacalcaactctcctttcct	53
stSG30448		EST		acactgcttgggggacaggct	43	agttgcagagacclagcctgtc	54
W18677		EST		cacgacgccacagagccagctc	44	tctgggagaggacggagctggc	55

3. Radiation Hybrid (RH) Mapping: Radiation hybrid mapping was performed against the Genebridge4 panel (Gyapay et al., 1996, *Hum. Mol. Genet.* 5:339-46) purchased from Research Genetics, in order to refine the chromosomal localization of genetic markers used in genotyping as well as to

identify, confirm, and refine localizations of markers from proprietary sequences. Standard PCR procedures were used for typing the RH panel with markers of interest. Briefly, 10 µl PCR reactions contained 25 ng DNA of each of the 93 Genebridge4 RH samples. PCR products were electrophoresed on  
5 2% agarose gels (Sigma) containing 0.5 µg/ml ethidium bromide in 1 X TBE at 150 volts for 45 min. Model A3-1 electrophoresis systems were used (Owl Scientific Products, Portsmouth, NH). Typically, gels contained 10 tiers of lanes with 50 wells/tier. Molecular weight markers (100 bp ladder, GibcoBRL, Rockville, MD) were loaded at both ends of the gel. Images of the gels were  
10 captured with a Kodak DC40 CCD camera and processed with Kodak 1D software (www.kodak.com). The gel data were exported as tab delimited text files; names of the files included information about the panel screened, the gel image files and the marker screened. These data were automatically imported using a customized Perl script into Filemaker databases for data storage and  
15 analysis. The data were then automatically formatted and submitted to an internal server for linkage analysis to create a radiation hybrid map using RHMAPPER (L. Stein et al., 1995; available from Whitehead Institute/MIT Center for Genome Research, at <http://www.genome.wi.mit.edu/ftp/pub/software/rhmapper/>, and via anonymous ftp to <ftp://ftp.genome.wi.mit.edu>,  
20 in the directory /pub/software/rhmapper.)

4. BAC Library Screening: The protocol used for BAC library screening was based on the "overgo" method, originally developed by John McPherson at Washington University in St. Louis (<http://www.tree.caltech.edu/protocols/overgo.html>, and W-W. Cai et al., 1998, *Genomics* **54**:387-397).  
25 This method involved filling in the overhangs generated after annealing two primers, each 22 nucleotides in length, which overlap by 8 nucleotides. The resulting labeled 36 bp product was then used in hybridization-based screening of high density grids derived from the RPCI-11 BAC library (deJong, *supra*). Typically, 15 probes were pooled together to hybridize 12 filters (13.5 genome  
30 equivalents).

Stock solutions (2 µM) of combined complementary oligos were heated

at 80°C for 5 min, placed at 37°C for 10 min, and then stored on ice. Labeling reactions included the following: 1.0 µl H<sub>2</sub>O; 5 µl mixed oligos (2 µM each); 0.5 µl BSA (2 mg/ml); 2 µl OLB (-A, -C, -N6) Solution (see below); 0.5 µl <sup>32</sup>P-dATP (3000 Ci/mmol); 0.5 µl <sup>32</sup>P-dCTP (3000 Ci/mmol); and 0.5 µl Klenow fragment (5 U/µl). The reaction was incubated at room temperature for 1 hr, and unincorporated nucleotides were removed using Sephadex G50 spin columns.

Solution O: 1.25 M Tris-HCL, pH 8, 125 M MgCl<sub>2</sub>; Solution A: 1 ml Solution O, 18 µl 2-mercaptoethanol, 5µl 0.1M dTTP, 5µl 0.1 M dGTP; Solution B: 2 M HEPES-NaOH, pH 6.6; Solution C: 3 mM Tris-HCl, pH 7.4, 0.2 mM EDTA; Solutions A, B, and C were combined to a final ratio of 1:2.5:1.5, and aliquots were stored at -20°C.

High-density BAC library membranes were pre-wetted in 2 X SSC at 58°C. Filters were then drained slightly and placed in hybridization solution (1% BSA; 1 mM EDTA, pH 8.0; 7% SDS; and 0.5 M sodium phosphate), pre-warmed to 58°C, and incubated at 58°C for 2-4 hr. Typically, 6 filters were hybridized in each container. Ten milliliters of pre-hybridization solution was removed, combined with the denatured overgo probes, and added back to the filters. Hybridization was performed overnight at 58°C. The hybridization solution was removed and filters were washed once in 2 X SSC, 0.1% SDS, followed by a 30 min wash in the same solution at 58°C. Filters were then washed in: 1) 1.5 X SSC and 0.1% SDS at 58°C for 30 min; 2) 0.5 X SSC and 0.1% SDS at 58°C for 30 min; and finally in 3) 0.1 X SSC and 0.1% SDS at 58°C for 30 min. Filters were then wrapped in Saran Wrap and exposed to film overnight. To remove bound probe, filters were treated in 0.1 X SSC and 0.1% SDS pre-warmed to 95°C and cooled room temperature. Clone addresses were determined as described by instructions supplied by RPCI.

To recover clonal BAC cultures from the library, a sample from the appropriate library well was plated by streaking onto LB agar (T. Maniatis et al., 1982, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) containing 12.5 µg/ml chloramphenicol (Sigma), and plates were incubated overnight. A single colony and a portion of the initial

streak quadrant were inoculated into 400 µl LB plus chloramphenicol in wells of a 96 well plate. Cultures were grown overnight at 37°C. For storage, 100 µl of 80% glycerol was added and the plates placed at -80°C.

To determine the marker content of clones, aliquots of the 96 well plate cultures were transferred to the surface of nylon filters (GeneScreen Plus, NEN) placed on LB/chloramphenicol Petri plates. Colonies were grown overnight at 37°C and colony lysis was performed by placing filters on pools of:  
1) 10% SDS for 3 min; 2) 0.5 N NaOH and 1.5 M NaCl for 5 min; and 3) 0.5 M Tris-HCl, pH 7.5, and 1 M NaCl for 5 min. Filters were then air-dried and washed free of debris in 2 X SSC for 1 hr. The filters were air-dried for at least 1 hr and DNA was crosslinked linked to the membrane using standard conditions. Probe hybridization and filter washing were performed as described above for the primary library screening. Confirmed clones were stored in LB containing 15% glycerol.

In certain cases, polymerase chain reaction (PCR) was used to confirm the marker content of clones. PCR conditions for each primer pair were initially optimized with respect to MgCl<sub>2</sub> concentration. The standard buffer was 10 mM Tris-HCl (pH 8.3), 50 mM KCl, MgCl<sub>2</sub>, 0.2 mM each dNTP, 0.2 µM each primer, 2.7 ng/µl human DNA, 0.25 units of AmpliTaq (Perkin Elmer) and MgCl<sub>2</sub> concentrations of 1.0 mM, 1.5 mM, 2.0 mM or 2.4 mM. Cycling conditions included an initial denaturation at 94°C for 2 min followed by 40 cycles at 94°C for 15 sec, 55°C for 25 sec, and 72°C for 25 sec followed by a final extension at 72°C for 3 min. Depending on the results from the initial round of optimization the conditions were further optimized if necessary. Variables included increasing the annealing temperature to 58°C or 60°C, increasing the cycle number to 42 and the annealing and extension times to 30 sec, and using AmpliTaqGold (Perkin Elmer).

5. BAC DNA Preparation: Several different types of DNA preparation methods were used for isolation of BAC DNA. The manual alkaline lysis miniprep protocol listed below (Maniatis et al., 1982) was successfully used for most applications, i.e., restriction mapping, CHEF gel

analysis and FISH mapping, but was not reproducibly successful in endsequencing. The Autogen protocol described below was used specifically for BAC DNA preparation for endsequencing.

For manual alkaline lysis BAC minipreps, bacteria were grown in 15 ml  
5 terrific broth (TB) containing 12.5 µg/ml chloramphenicol. Cultures were placed in a 50 ml conical tube at 37°C for 20 hr with shaking at 300 rpm. The cultures were centrifuged in a Sorvall RT 6000 D at 3000 rpm (1800 x g) at 4°C for 15 min. The supernatant was then aspirated as completely as possible. In some cases cell pellets were frozen at -20°C at this step for up to 2 weeks. The  
10 pellet was then vortexed to homogenize the cells and minimize clumping. Following this, 250 µl of P1 solution (50 mM glucose, 15 mM Tris-HCl, pH 8, 10 mM EDTA, and 100 µg/ml RNase A) was added and the mixture pipetted up and down to mix. The mixture was then transferred to a 2 ml Eppendorf tube. Subsequently, 350 µl of P2 solution (0.2 N NaOH, 1% SDS) was added,  
15 mixed gently, and the mixture was incubated for 5 min at room temperature. Then, 350 µl of P3 solution (3 M KOAc, pH 5.5) was added and mixed gently until a white precipitate formed. The solution was incubated on ice for 5 min and then centrifuged at 4°C in a microfuge for 10 min.

The supernatant was transferred carefully (avoiding the white  
20 precipitate) to a fresh 2 ml Eppendorf tube, and 0.9 ml of isopropanol was added; the solution was mixed and left on ice for 5 min. The samples were centrifuged for 10 min, and the supernatant removed carefully. Pellets were washed in 70% ethanol and air-dried for 5 min. Pellets were resuspended in 200 µl of TE8 (10 mM Tris-HCl, pH 8.0, 1.0 mM EDTA, pH 8.0), and RNase  
25 (Boehringer Mannheim, <http://biochem.boehringer-mannheim.com>) added to 100 µg/ml. Samples were incubated at 37°C for 30 min, then precipitated by addition of NH<sub>4</sub>OAc to 0.5 M and 2 volumes of ethanol. Samples were centrifuged for 10 min, and the pellets were washed with 70% ethanol. The pellets were air-dried and dissolved in 50 µl TE8. Typical yields for this DNA  
30 prep were 3-5 µg per 15 ml bacterial culture. Ten to 15 µl of DNA was used for *EcoRI* restriction analysis; 5 µl was used for *NotI* digestion and clone insert

sizing by CHEF gel electrophoresis.

Autogen 740 BAC DNA preparations for endsequencing were made by dispensing 3 ml of LB media containing 12.5 µg/ml of chloramphenicol into autoclaved Autogen tubes. A single tube was used for each clone. For  
5 inoculation, glycerol stocks were removed from -70°C storage and placed on dry ice. A small portion of the glycerol stock was removed from the original tube with a sterile toothpick and transferred into the Autogen tube. The toothpick was left in the Autogen tube for at least two min before discarding. After inoculation the tubes were covered with tape to ensure that the seal was  
10 tight. When all samples were inoculated, the tubes were transferred into an Autogen rack holder and placed into a rotary shaker. Cultures were incubated at 37°C for 16-17 hr at 250 rpm. Following this, standard conditions for BAC DNA preparation, as defined by the manufacturer, were used to program the Autogen. However, samples were not dissolved in TE8 as part of the program.  
15 DNA pellets were left dry.

When the program was completed, the tubes were removed from the output tray and 30 µl of sterile distilled and deionized H<sub>2</sub>O was added directly to the bottom of the tube. The tubes were then gently shaken for 2-5 sec and then covered with parafilm and incubated at room temperature for 1-3 hr. DNA  
20 samples were then transferred to an Eppendorf tube and used either directly for sequencing or stored at 4°C for later use.

6. BAC Clone Characterization: DNA samples prepared either by manual alkaline lysis or the Autogen protocol were digested with *EcoRI* for analysis of restriction fragment sizes. These data were used to compare the  
25 extent of overlap among clones. Typically 1-2 µg were used for each reaction. Reaction mixtures included: 1 X Buffer 2 (NEB); 0.1 mg/ml BSA (NEB); 50 µg/ml RNase A (Boehringer Mannheim); and 20 units of *EcoRI* (NEB) in a final volume of 25 µl. Digestions were incubated at 37°C for 4-6 hr. BAC DNA was also digested with *NotI* for estimation of insert size by CHEF gel analysis (see  
30 below). Reaction conditions were identical to those for *EcoRI*, except that 20 units of *NotI* were used. Six microliters of 6 X Ficoll loading buffer containing

bromphenol blue and xylene cyanol was added prior to electrophoresis.

*EcoRI* digests were analyzed on 0.6% agarose (Seakem, FMC Bioproducts, Rockland, ME) in 1X TBE containing 0.5 µg/ml ethidium bromide.

Gels (20 cm x 25 cm) were electrophoresed in a Model A4 electrophoresis unit (Owl Scientific) at 50 volts for 20-24 hr. Molecular weight size markers included undigested lambda DNA, *HindIII* digested lambda DNA, and *HaeIII* digested .X174 DNA. Molecular weight markers were heated at 65°C for 2 min prior to loading the gel. Images were captured with a Kodak DC40 CCD camera and analyzed with Kodak 1D software.

*NotI* digests were analyzed on a CHEF DRII (Bio-Rad) electrophoresis unit according to the manufacturer's recommendations. Briefly, 1% agarose gels (Bio-Rad pulsed field grade) were prepared in 0.5 X TBE, equilibrated for 30 min in the electrophoresis unit at 14 °C, and electrophoresed at 6 volts/cm for 14 hr with circulation. Switching times were ramped from 10 sec to 20 sec. Gels were stained after electrophoresis in 0.5 µg/ml ethidium bromide. Molecular weight markers included undigested lambda DNA, *HindIII* digested lambda DNA, lambda ladder PFG ladder, and low range PFG marker (all from NEB).

7. BAC Endsequencing: The sequence of BAC insert ends utilized DNA prepared by either of the two methods described above. The ends of BAC clones were sequenced for the purpose of filling gaps in the physical map and for gene discovery information. The following vector primers specific to the BAC vector pBACe3.6 were used to generate endsequence from BAC clones: pBAC 5'-2 (TGT AGG ACT ATA TTG CTC; SEQ ID NO:56) and pBAC 3'-1 (CGA CAT TTA GGT GAC ACT; SEQ ID NO:57).

The ABI dye-terminator sequencing protocol was used to set up sequencing reactions for 96 clones. The BigDye (ABI; PE Applied Biosystems) Terminator Ready Reaction Mix with AmpliTaq<sup>®</sup> FS, Part number 4303151, was used for sequencing with fluorescently labeled dideoxy nucleotides. A master sequencing mix was prepared for each primer reaction set including: 1600 µl of BigDye terminator mix (ABI; PE Applied Biosystems); 800 µl of 5 X CSA

buffer (ABI; PE Applied Biosystems); 800 µl of primer (either pBAC 5'-2 or pBAC 3'-1 at 3.2 µM). The sequencing cocktail was vortexed to ensure it was well-mixed and 32 µl was aliquoted into each PCR tube. Eight microliters of the Autogen DNA for each clone was transferred from the DNA source plate to a corresponding well of the PCR plate. The PCR plates were sealed tightly and centrifuged briefly to collect all the reagents. Cycling conditions were as follows: 1) 95°C for 5 min; 2) 95°C for 30 sec; 3) 50°C for 20 sec; 4) 65°C for 4 min; 5) steps 2 through 4 were repeated 74 times; and 6) samples were stored at 4°C.

At the end of the sequencing reaction, the plates were removed from the thermocycler and centrifuged briefly. Centri-Sep 96 plates were then used according to manufacturer's recommendations to remove unincorporated nucleotides, salts, and excess primers. Each sample was resuspended in 1.5 µl of loading dye, and 1.3 µl of the mixture was loaded on ABI 377 Fluorescent Sequencers. The resulting endsequences were then used to develop markers to rescreen the BAC library for filling gaps and were also analyzed by BLASTN searching for EST or gene content.

**EXAMPLE 5: Subcloning and Sequencing of BAC RPCI-11 1098L22**

The physical map of the chromosome 20 region provided the location of the BAC RPCI-11\_1098L22 clone that contains Gene 216 (see Figure 6). The BAC RPCI-11\_1098L22 clone was deposited as clone RP11-1098L22 with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 USA, under ATCC Designation No. PTA-3171, on March 14, 2001 according to the terms of the Budapest Treaty. DNA sequencing of BAC, RPCI-11-1098L22 from the region was completed. BAC RPCI-11-1098L22 DNA, (the "BAC DNA") was isolated according to one of two protocols: either a QIAGEN purification (QIAGEN, Inc., Valencia, CA, per manufacturer's instructions) or a manual purification using a method which was a modification of the standard alkaline lysis/Cesium Chloride preparation of plasmid DNA (see e.g., F.M. Ausubel et al., 1997, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY). Briefly, for the manual



protocol, cells were pelleted, resuspended in GTE (50 mM glucose, 25 mM Tris-Cl (pH 8), 10 mM EDTA) and lysozyme (50 mg/ml solution), followed by addition of NaOH/SDS (1% SDS and 0.2N NaOH) and then an ice-cold solution of 3M KOAc (pH 4.5-4.8). RnaseA was added to the filtered  
5 supernatant, followed by treatment with Proteinase K and 20% SDS. The DNA was then precipitated with isopropanol, dried, and resuspended in TE (10 mM Tris, 1 mM EDTA (pH 8.0)). The BAC DNA was further purified by cesium chloride density gradient centrifugation (Ausubel et al., 1997).

Following isolation, the BAC DNA was hydrodynamically sheared using  
10 HPLC (Hengen et al., 1997, *Trends in Biochem. Sci.*, **22**:273-274) to an insert size of 2000-3000 bp. After shearing, the DNA was concentrated and separated on a standard 1% agarose gel. A single fraction, corresponding to the approximate size, was excised from the gel and purified by electroelution (Sambrook et al., 1989).

15 The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The blunt-ended DNA was then ligated to unique *Bst*XI-linker adapters (5' GTCTTCACCCACGGGG (SEQ ID NO:58) and 5' GTGGTGAAGAC (SEQ ID NO:59) in 100-1000 fold molar excess). These linkers were complimentary to the *Bst*XI-cut pMPX vectors, while the overhang was not self-  
20 complimentary. Therefore, the linkers would not concatemerize, nor would the cut-vector re-ligate to itself easily. The linker-adapted inserts were separated from unincorporated linkers on a 1% agarose gel and purified using GeneClean (BIO 101, Inc., Vista, CA). The linker-adapted insert was then ligated to a modified pBlueScript vector to construct a "shotgun" subclone library. The  
25 vector contained an out-of-frame lacZ gene at the cloning site, which became in-frame in the event that an adapter-dimer was cloned. Such adapter-dimer clones gave rise to blue colonies, which were avoided.

All subsequent steps were based on sequencing by ABI377 automated DNA sequencing methods. Major modifications to the protocols are highlighted  
30 below. Briefly, the library was transformed into DH5-competent cells (GibcoBRL, DH5-transformation protocol). Transformed cells were plated onto

antibiotic plates containing ampicillin and IPTG/X-gal. The plates were incubated overnight at 37°C. White colonies were identified and then used to plate individual clones for sequencing. The cultures were grown overnight at 37°C. DNA was purified using a silica bead DNA preparation method (Ng et al., 1996, *Nucl. Acids Res.*, **24**:5045-5047). In this manner, 25 µg of DNA was obtained per clone.

These purified DNA samples were sequenced using ABI dye-terminator chemistry. The ABI dye terminator sequence reads were run on ABI377 machines and the data were directly transferred to UNIX machines following lane tracking of the gels. All reads were assembled using PHRAP (P. Green, *Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V*, Jan. 1996, p.157) with default parameters and quality scores. The assembly was done at 8-fold coverage and yielded 1 contig, BAC RPCI-11-1098L22. SEQ ID NO:5 (Figure 7) comprises a portion of the BAC that includes the genomic sequence of Gene 216.

#### **EXAMPLE 6: Gene Identification**

Any gene or EST mapping to the interval based on public map data or proprietary map data was considered a candidate respiratory disease gene. Public map data were derived from several sources: the Genome Database (GDB, <http://gdbwww.gdb.org/>), the Whitehead Institute Genome Center (<http://www-genome.wi.mit.edu/>), GeneMap98, UniGene, OMIM, dbSTS and dbEST (NCBI, <http://www.ncbi.nlm.nih.gov/>), the Sanger Centre (<http://www.sanger.ac.uk/>), and the Stanford Human Genome Center (<http://www-shgc.stanford.edu/>). Proprietary data was obtained from sequencing genomic DNA (cloned into BACs) or cDNAs (identified by direct selection, screening of cDNA libraries or full length sequencing of IMAGE Consortium (<http://www-bio.11nl.gov/bbrp/image.html>) cDNA clones).

1. Gene Identification from clustered DNA fragments. DNA sequences corresponding to gene fragments in public databases (GenBank and human dbEST) and proprietary cDNA sequences (IMAGE consortium and direct selected cDNAs) were masked for repetitive sequences and clustered

using the PANGEA Systems (Oakland, CA) EST clustering tool. The clustered sequences were then subjected to computational analysis to identify regions bearing similarity to known genes. This protocol included the following steps:

- a. The clustered sequences were compared to the publicly available  
5 UniGene database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997).

The parameters for this search were:  $E = 0.05$ ,  $v = 50$ ,  $B = 50$ , where  $E$  was the expected probability score cutoff,  $V$  was the number of database entries returned in the reporting of the results, and  $B$  was the number of sequence alignments returned in the reporting of the results (Altschul et al., 1990).

- 10 b. The clustered sequences were compared to the GenBank database (NCBI) using BLASTN2 (Altschul et al., 1997). The parameters for this search were  $E=0.05$ ,  $V=50$ ,  $B= 50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

- c. The clustered sequences were translated into protein sequences for all six reading frames, and the protein sequences were compared to a non-  
15 redundant protein database compiled from GenPept Swissprot PIR (NCBI). The parameters for this search were  $E = 0.05$ ,  $V = 50$ ,  $B = 50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

- d. The clustered sequences were compared to BAC sequences (see below) using BLASTN2 (Altschul et al., 1997). The parameters for this search  
20 were  $E=0.05$ ,  $V=50$ ,  $B=50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

2. Gene Identification from BAC Genomic Sequence: Following assembly of the BAC sequences into contigs, the contigs were subjected to computational analyses to identify coding regions and regions bearing DNA sequence similarity to known genes. This protocol included the following steps:

- 25 a. Contigs were degapped. The sequence contigs often contained symbols (denoted by a period symbol) that represented locations where the individual ABI sequence reads had insertions or deletions. Prior to automated computational analysis of the contigs, the periods were removed. The original data were maintained for future reference.

- 30 b. BAC vector sequences were "masked" within the sequence by using the program crossmatch (P. Green, <http://chimera.biotech.washington>).

edu\UWGC). Since the shotgun library construction detailed above left some BAC vector in the shotgun libraries, this program was used to compare the sequence of the BAC contigs to the BAC vector and to mask any vector sequence prior to subsequent steps. Masked sequences were marked by "X" in the sequence files, and remained inert during subsequent analyses.

c. *E. coli* sequences contaminating the BAC sequences were masked by comparing the BAC contigs to the entire *E. coli* DNA sequence.

d. Repetitive elements known to be common in the human genome were masked using CROSSMATCH (P. Green, University of Washington). In this implementation of crossmatch, the BAC sequence was compared to a database of human repetitive elements (J. Jerka, Genetic Information Research Institute, Palo Alto, CA). The masked repeats were marked by "X" and remained inert during subsequent analyses.

e. The location of exons within the sequence was predicted using the MZEF computer program (Zhang, 1997, *Proc. Natl. Acad. Sci.*, **94**:565-568) and GenScan gene prediction program (Burge and Karlin, *J. Mol. Biol.*, **268**:78-94).

f. The sequence was compared to the publicly available UniGene database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were:  $E = 0.05$ ,  $v = 50$ ,  $B = 50$ , where  $E$  was the expected probability score cutoff,  $V$  was the number of database entries returned in the reporting of the results, and  $B$  was the number of sequence alignments returned in the reporting of the results (Altschul et al., 1990).

g. The sequence was translated into protein sequences for all six reading frames, and the protein sequences were compared to a non-redundant protein database compiled from GenPept Swissprot PIR (NCBI). The parameters for this search were  $E = 0.05$ ,  $V = 50$ ,  $B = 50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

h. The BAC DNA sequence was compared to a database of clustered sequences using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were  $E=0.05$ ,  $V=50$ ,  $B=50$ , where  $E$ ,  $V$ , and  $B$  were defined as above. The database of clustered sequences was prepared utilizing

a proprietary clustering technology (PANGEA Systems, Inc.) using cDNA clones derived from direct selection experiments (described below), human dbEST sequences mapping to the 20p13-p12 region, proprietary cDNAs, GenBank genes, and IMAGE consortium cDNA clones.

5           i. The BAC sequence was compared to the sequences derived from the ends of BACs from the region on chromosomes 20 using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were  $E=0.05$ ,  $V=50$ ,  $B=50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

          j. The BAC sequence was compared to the GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this  
10           search were  $E = 0.05$ ,  $V = 50$ ,  $B = 50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

          k. The BAC sequence was compared to the STS division of GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The  
15           parameters for this search were  $E=0.05$ ,  $V=50$ ,  $B=50$ , where  $E$ ,  $V$ , and  $B$  were defined as above.

          l. The BAC sequence was compared to the Expressed Sequence Tag (EST) GenBank database (NCBI) using the BLASTN2 algorithm (Altschul et al., 1997). The parameters for this search were  $E=0.05$ ,  $V=50$ ,  $B=50$ , where  $E$ ,  
20            $V$ , and  $B$  were defined as above.

          c.     Mapping Analysis

          Through mapping analysis, BAC RPCI-11\_1098L22 (ATCC Designation No. PTA-3171) was identified as containing Gene 216. This BAC sequence (SEQ ID NO:5, Figure 7) included the genomic sequence of  
25           Gene 216 (SEQ ID NO:6; Figure 29), which corresponded to the cDNA sequence of Gene 216 (SEQ ID NO:1; Figure 24).

**EXAMPLE 7: Gene 216 cDNA Cloning and Expression Analysis**

          1.     Construction and screening of cDNA libraries: Directionally cloned cDNA libraries from normal lung and bronchial epithelium were  
30           constructed using standard methods (Soares et. al., 1994, *Automated DNA Sequencing and Analysis*, Adams et al. (eds), Academic Press, NY, pp. 110-

114). Total and cytoplasmic RNAs were extracted from tissue or cells by homogenizing the sample in the presence of Guanidinium Thiocyanate-Phenol-Chloroform extraction buffer (e.g. Chomczynski and Sacchi, 1987, *Anal. Biochem.*, **162**:156-159) using a polytron homogenizer (Brinkman Instruments, <http://www.brinkmann.com>). Poly A + RNA was isolated from total/cytoplasmic RNA using dynabeads-dT according to the manufacturer's recommendations (DynaI, Inc., <http://www.dynal.com>). The double stranded cDNA was then ligated into the plasmid vector pBluescript II KS+ (Stratagene, <http://www.stratagene.com>), and the ligation mixture was transformed into *E. coli* host DH10B or DH12S by electroporation (Soares, 1994). Following overnight growth at 37°C, DNA was recovered from the *E. coli* colonies after scraping the plates by processing as directed for the Mega-prep kit (QIAGEN). The quality of the cDNA libraries was estimated by counting a portion of the total number of primary transformants, determining the average insert size, and the percentage of plasmids with no cDNA insert. Additional cDNA libraries (human total brain, heart, kidney, leukocyte, and fetal brain) were purchased from Life Technologies (Bethesda, MD).

cDNA libraries, both oligo (dT) and random hexamer-primed, were used for isolating cDNA clones mapped within the disorder critical region. Four 10 x 10 arrays of each of the cDNA libraries were prepared as follows. The cDNA libraries were titered to  $2.5 \times 10^6$  using primary transformants. The appropriate volume of frozen stock was used to inoculate 2 L of LB/ampicillin (100 µg/µl). Four hundred aliquots containing 4 ml of the inoculated liquid culture were generated. Each tube contained about 5000 cfu (colony forming units). The tubes were incubated at 30°C overnight with shaking until an OD of 0.7-0.9 was obtained. Frozen stocks were prepared for each of the cultures by aliquotting 300 µl of culture and 100 µl of 80% glycerol. Stocks were frozen in a dry ice/ethanol bath and stored at -70°C. DNA was isolated from the remaining culture using the QIAGEN spin mini-prep kit according to the manufacturer's instructions. The DNA from the 400

cultures were pooled to make 80 column and row pools. Markers were designed to amplify putative exons from candidate genes. Once a standard PCR condition was identified and specific cDNA libraries were determined to contain cDNA clones of interest, the markers were used to screen the  
5 arrayed library. Positive addresses indicating the presence of cDNA clones were confirmed by a second PCR using the same markers.

Once a cDNA library was identified as likely to contain cDNA clones corresponding to a transcript of interest from the disorder critical region, it was used to isolate a clone or clones containing cDNA inserts. This was  
10 accomplished by a modification of the standard "colony screening" method (Sambrook et al., 1989). Specifically, twenty 150 mm LB plus ampicillin agar plates were spread with 20,000 cfu of cDNA library. Colonies were allowed to grow overnight at 37°C. Colonies were then transferred to nylon filters (Hybond from Amersham-Pharmacia, or equivalent) and duplicates prepared  
15 by pressing two filters together essentially as described (Sambrook et al., 1989). The "master" plate was then incubated an additional 6-8 hr to allow the colonies additional growth. The DNA from the bacterial colonies was then bound to the nylon filters by treating the filters sequentially with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for 2 min, and neutralization  
20 solution (0.5 M Tris-Cl pH 8.0, 1.5 M NaCl) for 2 min (twice). The bacterial colonies were removed from the filters by washing in a solution of 2 X SSC/2% SDS for 1 min while rubbing with tissue paper. The filters were air-dried and baked under vacuum at 80°C for 1-2 hr to crosslink the DNA to the filters.

25 cDNA hybridization probes were prepared by random hexamer labeling (Fineberg and Vogelstein, 1983, *Anal. Biochem.*, **132**:6-13) or by including gene-specific primers and no random hexamers in the reaction (for small fragments). The colony membranes were then pre-washed in 10 mM Tris-Cl pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS for 30 min at 55°C.  
30 Following the pre-wash, the filters were pre-hybridized in > 2 ml/filter of 6 X SSC, 50 % deionized formamide, 2% SDS, 5 X Denhardt's solution, and 100

mg/ml denatured salmon sperm DNA, at 42°C for 30 min. The filters were then transferred to hybridization solution (6 X SSC, 2% SDS, 5 X Denhardt's, 100 mg/ml denatured salmon sperm DNA) containing denatured  $\alpha$ -<sup>32</sup>P-dCTP-labeled cDNA probe and incubated overnight at 42°C.

5           The following morning, the filters were washed under constant agitation in 2 X SSC, 2% SDS at room temperature for 20 min, followed by two washes at 65°C for 15 min each. A second wash was performed in 0.5 X SSC, 0.5% SDS for 15 min at 65°C. Filters were then wrapped in plastic wrap and exposed to radiographic film. Individual colonies on plates were  
10 aligned with the autoradiograph and positive clones picked into a 1 ml solution of LB Broth containing ampicillin. After shaking at 37°C for 1-2 hr, aliquots of the solution were plated on 150 mm plates for secondary screening. Secondary screening was identical to primary screening (above) except that it was performed on plates containing ~250 colonies so that  
15 individual colonies could be clearly identified. Positive cDNA clones were characterized by restriction endonuclease cleavage, PCR, and direct sequencing to confirm the sequence identity between the original probe and the isolated clone.

          To obtain the full-length cDNA, novel sequence from the 5'-end of the  
20 clone was used to reprobe the library. This process was repeated until the length of the cDNA cloned matched that of the mRNA, estimated by Northern analysis. Utilizing this process, a single uterus clone was isolated and deposited as clone Gene 216\_CS759 with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 USA,  
25 under ATCC Designation No. PTA-3173, on March 14, 2001, according to the terms of the Budapest Treaty. The uterus clone (SEQ ID NO:3) contained the entire Gene 216 open reading frame. Both strands of this clone were completely sequenced and the data were compared against the BAC sequence. Any discrepancies were flagged, and these regions were  
30 resequenced. The final analysis of the sequence revealed that the uterine clone was 3433 bp long and contained the full complement of exons defining



the open reading frame (SEQ ID NO:3). In addition, the clone contained a small portion of the 5' untranslated region (5 bp), the entire 3' untranslated region with a polyadenylation signal, and a poly A tail of 76 bp in length. The Gene 216 open reading frame was determined to be 2436 bp in length and to encode a protein of 812 amino acids (SEQ ID NO:363). Analysis of the composition of SNPs across the cDNA clone revealed that it contained the most frequent haplotype (Figure 8, see below).

Rapid Amplification of cDNA ends (RACE) was performed following the manufacturer's instructions using a Marathon cDNA Amplification Kit (CLONTECH) as a method for cloning the 5' and 3' ends of candidate genes. cDNA pools were prepared from total RNA by performing first strand synthesis. For first strand synthesis, a sample of total RNA sample was mixed with a modified oligo (dT) primer, heated to 70°C, cooled on ice and incubated with: 5 X first strand buffer (CLONTECH), 10 mM dNTP mix, and AMV Reverse Transcriptase (20 U/μl). The reaction mixture was incubated at 42°C for 1 hr and placed on ice. For second-strand synthesis, the following components were added directly to the reaction tube: 5 X second-strand buffer (CLONTECH), 10 mM dNTP mix, sterile water, and 20 X second-strand enzyme cocktail (CLONTECH). The reaction mixture was incubated at 16°C for 1.5 hr. T4 DNA Polymerase was added to the reaction mixture and incubated at 16°C for 45 min. The second-strand synthesis was terminated with the addition of an EDTA/Glycogen mix. The sample was purified by phenol/chloroform extraction and ammonium acetate precipitation. The cDNA pools were checked for quality by analyzing on an agarose gel for size distribution. Marathon cDNA adapters were then ligated onto the cDNA ends. The specific adapters contained priming sites that allowed for amplification of either 5' or 3' ends, and varied depending on the orientation of the gene specific primer (GSP) that was chosen. An aliquot of the double stranded cDNA was added to the following reagents: 10 μM Marathon cDNA adapter, 5 X DNA ligation buffer, T4 DNA ligase. The reaction was incubated at 16°C overnight and heat inactivated to terminate

the reaction. PCR was performed by the addition of the following to the diluted double stranded cDNA pool: 10X cDNA PCR reaction buffer, 10  $\mu$ M dNTP mix, 10  $\mu$ M GSP, 10  $\mu$ M AP1 primer (kit), 50 X Advantage cDNA Polymerase Mix. Thermal Cycling conditions were carried out at 94°C for 30 sec; 5 cycles of 94°C for 5 sec, 72°C for 4 min, 5 cycles of 94°C for 5 sec, and 70°C for 4 min; 23 cycles of 94°C for 5 sec; 68°C for 4 min. The first round of PCR was performed using the GSP to extend to the end of the adapter to create the adapter primer-binding site. Following this, exponential amplification of the specific cDNA of interest was performed. Usually, a second, nested PCR was performed to provide specificity. The RACE product was analyzed on an agarose gel. Following excision from the gel and purification (GeneClean, BIO 101), the RACE product was then cloned into pCTNR (General Contractor DNA Cloning System, 5' - 3', Inc.) and sequenced to verify that the clone was specific to the gene of interest. The 5' RACE technique was employed to identify the 5' untranslated region of Gene 216. Experiments were performed using lung mRNA and a primer that hybridized near the 5' end of the available sequence. The result of the experiment identified an additional 75 bp 5' of that present in the uterus cDNA clone (rt690; SEQ ID NO:351). This sequence was subsequently cloned and deposited with the ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209 USA), as clone Gene 216\_rt690, under ATCC Designation No.PTA-3172 on March 14, 2001, according to the terms of the Budapest Treaty.

Further attempts to extend the 5' end of Gene 216 by 5' RACE gave similar results indicating that the 5' end of the transcript was obtained.

This sequence in combination with the uterus cDNA clone yielded the master consensus sequence containing the 5' to 3' cDNA for Gene 216 (SEQ ID NO:1; Figure 24).

2. Identification of Splice Variants: Additional cDNA clones were isolated that represented alternatively spliced variants of Gene 216. To ensure that all splice variants present in lung tissue were identified, an RT-PCR-based

screening protocol was designed using multiple primer pairs spanning the entire gene. These amplicons produced PCR fragments of approximately 600 bp and overlapped by approximately 100 bp. The PCR products were fractionated on agarose gels and any fragments that were different from the expected size were cloned and sequenced. These results are summarized in Figures 9 and 10. The availability of the complete genomic sequence of BAC RPCI-11\_1098L22 enabled the intron/exon structure of Gene 216 (Figure 11) to be determined. Gene 216 contains 21 exons that span approximately 23.5 kb of genomic DNA.

Analysis of the sequence surrounding the intron/exon boundaries indicated that the consensus splice sequence GT/AG was upheld in all cases (Table 4). However, in several of the cDNA clones, an alternative use of a splice site at the intron/exon boundary of exon T was identified. The sequence CAGCAG was present at the border of intron ST and exon T resulting in a duplication of the canonical acceptor splice consensus CAG. Typically, a C residue preceding the AG is found in approximately 65% of acceptor splice sites. As a consequence, the splicing machinery can utilize either AG resulting in the presence or absence of an alanine. If the first AG (splice site 1) were utilized near the junction of intron ST and exon T, the resulting protein would encode the amino acid sequence DPQADQVQM (Figure 12) (SEQ ID NO:60). However, if the second AG (splice site 2) were favored, then one alanine would be omitted from the amino acid sequence and the protein would contain the amino acid sequence DPQDQVQM (Figure 12) (SEQ ID NO:61). The percentage that used splice site 1 or splice site 2 could not be determined from the dataset because the majority of the clones were derived from PCR-based techniques.

**TABLE 4**

EXON	3' INTRON	5' EXON	3' EXON	5' INTRON
A			AAG	GTGAGG
B	CAG	GAC	CCG	GTCAAGT
C	CAG	GTC	CCA	GTGAGT
D	CAG	CAG	ACG	GTGAGA
D (ALT)	CAG	CAG	GAG	GTACCC
E	TAG	GAT	GAG	GTGAGC
F	TAG	TGG	AGG	GTCAAGG
G	CAG	GGC	CTG	GTGAGG
H	CAG	TTC	CAG	GTTGGG
I	CAG	CTT	CAC	GTGGGT
J	CAG	GGG	ACG	GTGAGC
K	CAG	GAC	CGG	GTACGC
L	TAG	GCA	CAG	GTTAAG
L2	CAG	GAG	CTG	GTGAGG
M	CAG	CTG	CTG	GTGAGA
N	CAG	GCT	GAG	GTAGGG
O	CAG	GGA	ATG	GTGAGC
O (ALT)	TAG	ATG	ATG	GTGAGC
U	TAG	GTG	GGG	GTGAGA
P	CAG	GTT	AAA	GTATGC
Q	CAG	ACC	TGG	GTAGGC
R	CAG	CCC	TGG	GTGAGT
S	CAG	ACC	AAG	GTAGGC
T	CAG	CAG		

C<sub>65</sub> A<sub>100</sub> G<sub>100</sub> N A<sub>64</sub> G<sub>73</sub> G<sub>100</sub> T<sub>100</sub> A<sub>62</sub> A<sub>68</sub> G<sub>84</sub> T<sub>63</sub>

3. Promoter Analysis: In order to identify the transcriptional start site of Gene 216, multiple 5' RACE products were sequenced from several different tissues. In most cases the 5' ends were located 80 bp upstream of the translational start site. The region upstream of this sequence was then analyzed for potential transcription factor binding sites using GEMS Launcher, a promoter analysis program (<http://anthea.gsf.de/>). GEMS Launcher uses statistically weighted algorithms to identify binding elements that comprise a promoter or regulatory module. A stretch of DNA sequence spanning the 2000 bp upstream of the translational start site was analyzed. The results indicated that Gene 216 did not possess a TATA or CCAAT box. In fact, the first binding element that was identified was a GC box within the 5' untranslated region oriented in the opposite direction (Figure 13). This result is not unprecedented since 60% of TATA-less genes possess a GC box on the opposing strand. Also, this result was in agreement with published data regarding the promoters of mouse ADAM 17 and 19. Other binding elements that were identified within

600 bp upstream of the initiator methionine included an E-box, one AP2, and three SP1 sites (Figure 13). These types of binding elements were also identified in the mouse ADAM 17 and 19 genes and may represent components of a promoter module for Gene 216. Approximately 1200 bp  
5 upstream of the putative promoter module, GEMS Launcher identified binding elements that may comprise an additional regulatory element (Figure 13). This region was highly conserved with the mouse ortholog of Gene 216 (see below), as determined by dot matrix analysis.

4. BLAST Analysis: BLASTP, BLASTN, and BLASTX analysis of  
10 Gene 216 against protein and nucleotide databases revealed that it was a novel member of the ADAM (A Disintegrin And Metalloprotease) gene family. This gene family, of which there are currently 31 members, is a sub-group of the zinc-dependent metalloprotease superfamily. ADAMs have a complex domain organization that includes a signal sequence, propeptide,  
15 metalloprotease, disintegrin, cysteine-rich, and epidermal growth factor-like domains, as well as a transmembrane region and cytoplasmic tail. ADAM proteins have been implicated in many processes such as proteolysis in the secretory pathway and extracellular matrix, extra- and intra-cellular signaling, processing of plasma membrane proteins and procytokine conversion. The  
20 homology of Gene 216 and human ADAMs 19, 12, 15, 8 and 9 indicated that Gene 216 belonged to a branch of the 31-member family containing active metalloprotease domains (Figure 14).

6. Expression Analysis: To characterize the expression of Gene 216, a series of expression experiments were performed.

25 i. Northern Analysis: To characterize novel genes, Northern analysis (Sambrook et al., 1989) can be used to determine the length, in nucleotides, of the processed transcript or messenger RNA (mRNA). Probes were generated using one of the methods described below. Briefly, sequence verified IMAGE consortium cDNA clones were digested with appropriate  
30 restriction endonucleases to release the insert. The restriction digest was electrophoresed on an agarose gel and the bands containing the insert were

excised. The gel piece containing the DNA insert was placed in a Spin-X (Corning Costar Corporation, Cambridge, MA) or Supelco spin column (Supelco Park, PA) and spun at high speed for 15 min. The DNA was ethanol precipitated and resuspended in TE. Alternatively, PCR products obtained from genomic DNA or RT-PCR were purified. First, oligonucleotide primers were designed for use in the polymerase chain reaction (PCR) so that portions of the cDNA, EST, or genomic DNA could be amplified from a pool of DNA molecules or RNA population (RT-PCR). The PCR primers were used in a reaction containing genomic DNA to verify that they generated a product of the predicted size (based on the genomic sequence. Inserts purified from IMAGE clones or PCR products were random primer labeled (Fineberg and Vogelstein, *supra*) to generate probes for hybridization. Probes from purified PCR products were generated by incorporation of  $\alpha$ -<sup>32</sup>P-dCTP in second round of PCR. Commercially available Multiple Tissue Northern blots (CLONTECH) were hybridized and washed under conditions recommended by the manufacturer. A separate filter that contained 6 tissues from the immune system was also utilized. The results revealed a major 5.0 kb transcript and a minor 3.5 kb transcript that were expressed in most tissues examined (Figures 15A-15B). The strongest signals were consistently identified in heart, skeletal muscle, colon, lymph, and small intestine, with lung, liver, kidney, placenta, bone marrow, and brain showing moderate expression levels.

The 5 kb transcript was further analyzed to determine if it was an incompletely spliced version of the Gene 216 transcript. To test this hypothesis, Northern blotting was performed using cytoplasmic mRNA isolated from bronchial smooth muscle cells. The same radioactive probe was employed as previously. The results showed a very strong 3.5 kb signal and no signal at 5.0 kb (Figure 15C) suggesting that the predominant 5 kb transcript contained intronic material and was localized to the nucleus. Interestingly, intron QR is 1.4 kb in size. The addition of the QR intron and the 3.5 kb full length cDNA would total ~5.0 kb. Accordingly, there may be regulatory elements within the region around intron QR that affect splicing,

retention in the nucleus, and/or transport to the cytoplasm.

ii. RNA Dot Blot Analysis: RNA dot blotting was used to determine the expression of Gene 216 in a wide range of tissues. mRNA from 50 tissues was dotted onto a nylon filter, and a radioactive probe  
5 designed to hybridize to the 3' untranslated region was used. Figure 16 shows that Gene 216 was highly expressed in gastrointestinal tissues as well as aorta, uterus, prostate, ovary, lung, fetal lung, trachea and placenta.

Notably, the majority of these tissues are derived from the endoderm, which forms a tube that produces the primordium of the digestive tract. Extensions  
10 from this wall also develop into organs such as the lung and trachea.

iii. RT-PCR: Total RNA isolated from primary cultures of seven cell types cultured from lung tissue was analyzed in RT-PCR experiments. Genomic DNA was removed from the total RNA by DNaseI digestion. The "Superscript" Preamplification System for First strand cDNA synthesis" (Life  
15 Technologies) was used according to manufacturer's specifications with oligo(dT) or random hexamers to synthesize cDNA from the DNaseI treated total RNA. Gene specific primers were used to amplify the target cDNAs in a 30 µl PCR reaction containing 0.5 µl of first strand cDNA, 1 µl sense primer (10 µM), 1 µl antisense primer (10 µM), 3 µl dNTPs (2 mM), 1.2 µl MgCl<sub>2</sub> (25 mM),  
20 3 µl 10 X PCR buffer and 1 unit of Taq Polymerase (Perkin Elmer). The PCR reaction was initially incubated at 94°C for 4 min, followed by 30 cycles of incubation at 94°C for 30 sec, 58°C for 1 min, and 72°C for 1 min; then followed by a final incubation at 72°C for 7 min. PCR products were analyzed on agarose gels. Figure 17 shows that Gene 216 was expressed in lung  
25 fibroblasts, pulmonary artery smooth muscle cells, bronchial smooth muscle cells and total lung, but not in bronchial epithelium or pulmonary artery endothelial cells.

iv. cDNA Library Representation: A comprehensive approach to determining the tissue distribution of Gene 216 was performed *in silico* by  
30 mining the public EST database and Genome Therapeutics Corporation's internal cDNA database. BLAST analysis identified ESTs from multiple cDNA

libraries. A summary of all tissues expressing Gene 216 is given in Table 5.

**TABLE 5**

Source	Tissue
UNIGENE	Eye Muscle Placenta Stomach Uterus Whole embryo Breast Normal testis
Direct selected cDNAs	Bronchial smooth muscle (1 clone) Normal lung (2 clones) Brain (1 clone)
Primary cell types (RT/PCR)	Pulmonary artery smooth muscle Bronchial smooth muscle Lung fibroblast Total lung
RNA Dot Blot	Aorta Colon Bladder Uterus Prostate Ovary Small intestine Heart Stomach Testis Appendix Lung Trachea Fetal kidney Fetal lung
Northern Blot	Brain Heart Skeletal muscle Colon Thymus Spleen Kidney Liver Small intestine Placenta Lung Lymph Bone marrow



**EXAMPLE 8: Gene 216 Polypeptide**

1. ADAM Family Features: The zinc-dependent metalloprotease superfamily is comprised of several sub-groups. Those proteases that exhibit the characteristic Zn-binding consensus sequence HEXXHXXGXXH (SEQ ID NO:62) are referred to as zincins. The 3 histidines play an essential role in binding to the catalytically essential zinc ion. The zincins can be further classified into metzincins if a methionine residue is located beneath the active-site zinc ion ("Met-turn" motif). Within this sub-group there are 4 sub-families: astacins, matraxins, adamlysins, and serralsins. The ADAM genes fall within the adamlysins sub-family along with snake venom metalloproteases.

Currently, there are 31 members of the ADAM family. The ADAM genes encode proteins of approximately 750 amino acids with 8 different domains. Domain I is a pre-domain and contains the signal sequence peptide that facilitates secretion through the plasma membrane. Domain II is a pro-domain that is cleaved before the protein is secreted resulting in activation of the catalytic domain. Domain III is a catalytic domain containing metalloprotease activity. Domain IV is a disintegrin-like domain and is believed to interact with integrins or other receptors. Domain V is a cysteine-rich domain and is speculated to be involved in protein-protein interactions or in the presentation of the disintegrin-like domain. Domain VI is an EGF-like domain that plays a role in stimulating membrane fusion. Domain VII is a transmembrane domain that anchors the ADAM protein to the membrane. Domain VIII is a cytoplasmic domain and contains binding sites for cytoskeletal-associated proteins and/or SH3 binding domains that may play a role in bi-directional signaling. See Figure 8 for the location of ADAM domains identified in the Gene 216 protein sequence.

To determine whether Gene 216 was a novel member of the ADAM family, the 812 amino acid sequence was aligned by Pile-Up (Genetics Computer Group, <http://www.gcg.com>) (Figure 18). These analyses indicated that Gene 216 possessed the characteristic consensus sequence HEXXHXXGXXH (SEQ ID NO:62) located within the catalytic domain. In

addition, a methionine residue referred to as a "Met-turn" was identified in the Gene 216 protein. A conserved cysteine (amino acid 133 in Gene 216) that plays a role in activating ADAM proteins was identified in the prodomain of Gene 216 protein. In ADAM proteins, this single cysteine residue forms an intramolecular complex with the zinc ion bound to the metalloprotease domain and blocks the active site. The catalytic domain is activated by the dissociation of the cysteine from the complex, resulting in either a conformational change or enzymatic cleavage of the prodomain. This process is referred to as the "cysteine switch".

10 In ADAM 12, the position of the cysteine residue was reported to be located in a different position in the prodomain (B.L. Gilpin et al., 1998, *J. Biol. Chem.* **273**:157-166). This location would correspond to the cysteine residue at amino acid 179 in Gene 216 (Figure 19). However, in accordance with analyses performed by Stone et al., using 14 ADAMs, including ADAMs 8, 9, 15 12 and 15, the cysteine residue corresponding to position 133 of Gene 216 (Figures 18 and 19) was identified as being involved in the "cysteine switch". In addition, there appeared to be more sequence identity around the cysteine at amino acid 133 in Gene 216 than at position 179. This provided further support that the cysteine at position 133 was involved in the "cysteine switch". 20 The alignment also indicated that the amino acid sequence of Gene 216 contained all eight domains that define the hallmarks of these types of genes (Figure 18).

Hydrophobicity analysis (PepPlot, Genetics Computer Group) of the Gene 216 amino acid sequence revealed the presence of two hydrophobic regions (Figure 20). One region is located at the amino terminus of the protein and is the putative the signal sequence. The other hydrophobic region is located near the carboxyl terminus and is the putative transmembrane domain that anchors the protein to the cell surface. Computational biology analysis (<http://blocks.fhcrc.org>) of the Gene 216 cytoplasmic domain revealed the presence of a putative SH2 and SH3 binding domain as well as a putative 30

casein kinase I phosphorylation site (Figure 19). These sites may contribute to a role in bi-directional signaling, a function attributed to ADAM proteins.

Sequence analyses indicated that Gene 216 is a novel member of the ADAM family. Gene 216 is most closely related to ADAMs 8, 9, 12, 15, and 19, a branch of the family that is known to possess an active metalloprotease domain. Table 6 lists the 5 most similar BLASTP hits using the Gene 216 amino acid sequence as a query. Based on BLASTN and BLASTP analysis, Gene 216 nucleotide sequence shares the 37% identity with the ADAM 19 nucleotide sequence; and Gene 216 amino acid sequence shares 58% identity with the ADAM 19 amino acid sequence.

**Table 6: Top 5 Hits from BLAST Analysis of Gene 216 protein**

Hit	GenBank Locus	Description	Smallest Sum
1	U66003	<i>Xenopus laevis</i> (ADAM 13)	5.5e-166
2	AF019887	<i>Mus musculus</i> metalloprotease-disintegrin meltrin beta	1.2e-139
3	AF134707	<i>Homo sapiens</i> disintegrin and metalloprotease domain 19 (ADAM19)	1.6e-139
4	S60257	Mouse mRNA for meltrin alpha	1.8e-121
5	AF023476	<i>Homo sapiens</i> meltrin-L precursor (ADAM12)	4.9e-119

Table 7 lists the top two hits from BLIMPS analysis of the Block protein motif database (<http://blocks.fhcrc.org/>).

**Table 7: Top 2 Hits from BLIMPS Analysis of Gene 216 protein**

Description	Strength	Score	AA#	AA	Sequence
Disintegrins proteins	1950	1597	377		CCfAhnCsLRPGAQCAh- GdCCvRCIIKpAGal- CRqAMGDCDIPEfCT- GTSshCPP (SEQ ID NO:335)
Zinc metallopeptidases	1173	1276	276		TMAHEIGHSLG (SEQ ID NO:336)

2. Amino Acid Changes: In total, there were 9 SNPs within the open reading frame of Gene 216. See Example 10 for details on polymorphism identification and Figure 19 for resulting changes to the protein sequence.

- 5 Seven of the nine SNPs constituted an amino acid change and the other 2 were synonymous. Of the 7 amino acid changes, 4 were clustered toward the carboxyl terminus of the protein: one within the identified transmembrane domain and 3 within the identified cytoplasmic domain.

One SNP located in an identified SH2 binding domain resulted in a  
10 significant amino acid change: methionine (hydrophobic) to threonine (polar). The remaining two SNPs in the identified cytoplasmic domain resulted in significant amino acid changes: proline (hydrophobic) to serine (polar) and glutamine (polar) to histidine (basic). These amino acid changes may disturb the signaling properties of the Gene 216 protein. In addition, the valine to  
15 isoleucine amino acid change in the putative transmembrane domain may affect signaling efficiency.

The two SNPs in the identified pro-domain generated significant amino acid changes: tyrosine (polar) to histidine (basic) and threonine (polar) to alanine (hydrophobic). Since the ADAM pro-domain is cleaved during  
20 activation of the catalytic domain, it is possible that these amino acid changes affect the cleavage process. One SNP in the identified catalytic domain resulted in a change from alanine (hydrophobic) to valine (hydrophobic). This amino acid change may affect sheddase efficiency.

Notably, amino acid changes in the identified Gene 216 catalytic  
25 domain, especially within the metalloprotease domain, would be of great interest, as this domain is critical to sheddase function. Recently, the X-ray crystallographic data of the snake venom catalytic domain was determined and deposited in the public domain (<http://www.rcsb.org/pdb/cgi/explore.cgi?pid=9267984771616&pdbld=1C9G>; Accession No. 1C9GA). This information  
30 can be utilized to determine whether an amino acid change alters the folding of the catalytic domain of the Gene 216 protein. In particular, the sequence of

the catalytic domain of Gene 216 protein can be plotted as X-ray crystallographic coordinates and used to determine changes in the tertiary structure of this domain.

3. Biological Role of Gene 216: ADAMs are part of a very large  
5 superfamily called zinc-dependent metalloproteases (Stone et. al., 1999, *J. Prot. Chem.* **18**:447-465). Gene 216 represents a novel member of the ADAM family that is closely related to ADAM 19, a gene that was found to participate in the proteolytic processing of the membrane anchored protein neuregulin 1 (NRG1) (Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). The  
10 expression and activation of ADAM 19 protein has been localized to the trans-Golgi apparatus. This has been observed for other ADAM proteins (Lum et al., 1998, *J. Biol. Chem.* **273**:26236-26247; Roghani et. al., 1999, *J. Biol. Chem.* **274**:3531-3540; Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). These data suggest that the ADAM genes, and Gene 216, encode proteins  
15 that function in the trans-Golgi apparatus as intracellular processing enzymes. The processed substrates of these enzymes may be released into the cytosol as part of a signal transduction cascade leading to the cell surface.

The substrate of ADAM 19, NRG1, belongs to a group of growth factors (neuregulins) that are members of the epidermal growth factor family. The  
20 neuregulins participate in an array of biological effects that are mediated by the epidermal growth factor family of tyrosine kinase receptors. Data suggest that the proteolytically cleaved isoform of NRG1, NRG- $\beta$ 1, may induce the tyrosine phosphorylation of EGFR2 and EGFR3 in differentiated muscle cells (Shirakabe et. al., 2001, *J. Biol. Chem.* **276**(12):9352-8). The sequence  
25 similarity of Gene 216 protein and ADAM 19 protein suggests that the neuregulins or their isoforms serve as substrates for Gene 216 protein. The Gene 216-processed neuregulins or isoforms may then serve as ligands for EGFR1.

Epidermal growth factor receptor (EGFR1) plays a pivotal role in the  
30 maintenancé and repair of epithelial tissue. Following injury in bronchial epithelium, EGFR1 is upregulated in response to ligands acting on it or through

transactivation of the EGFR1 receptor. This results in the increased proliferation of cells and airway remodeling at the point of insult, leading to the repair of the bronchial epithelium (Polosa et. al., 1999, *Am. J. Respir. Cell Mol. Biol.* **20**:914-923; Holgate et. al., 1999, *Clin. Exp. Allergy Suppl* **2**:90-95).

5           In asthma, the bronchial epithelium is highly abnormal, with structural changes involving separation of columnar cells from their basal attachments and functional changes that include increased expression and release of proinflammatory cytokines, growth factors, and mediator-generating enzymes. Beneath this damaged structure are the subepithelial myofibroblasts that have  
10           been activated to proliferate. This, in turn, causes excessive matrix deposition leading to abnormal thickening and increased density of the subepithelial basement membrane.

          Immunocytochemical studies have shown that both TGF-  $\beta$  and EGFR1 are highly expressed at the area of injury and that parallel pathways could be  
15           operating in the repairing epithelial cells (Puddicombe et. al., 2000, *FASEB J.* **14**:1362-1374). EGFR1 stimulates epithelial repair and TGF-  $\beta$  regulates the production of profibrogenic growth factors and proinflammatory cytokines leading to extracellular matrix synthesis. As EGFR1 is involved in regulating  
20           a number of different stages of epithelial repair (survival, migration, proliferation and differentiation), any inhibitory effects that act on the receptor may cause the epithelium to be held in a "state of repair" (Holgate et. al., 1999, *Clin. Exp. Allergy Suppl* **2**:90-95).

          Without wishing to be bound by theory, it is possible that a variant Gene  
216 protein induces the epithelium into a continuous "state of repair" by  
25           functioning improperly and failing to release its substrate (a member of the neuregulin family) that serves as the ligand for EGFR1. This, in turn, may cause the observed increase in EGFR1 expression. Under these circumstances, the TGF-  $\beta$  pathway remains active, producing a continuous source of proinflammatory products as well as growth factors that drive airway  
30           wall remodeling causing bronchial hyperresponsiveness, a phenotype of asthma.

It is also possible that the disintegrin-like domain of Gene 216 plays a role in respiratory diseases. Integrins are a family of heterodimeric transmembrane receptors that mediate cell-cell and cell-extracellular matrix interaction (Hynes, 1992, *Cell* **69**:11). Integrins mediate angiogenesis (Brooks et al., 1994, *Science* **264**:569), which plays a major role in various pathological mechanisms, such as tumor growth, metastasis, diabetic retinopathy, and certain inflammation diseases (Folkman, 1995, *N. Engl. J. Med.* **333**:1757). Disintegrins act as integrin ligands that disrupt cell-matrix interactions (C.P. Blobel and J.M. White, 1992, *Curr. Opin. Cell Biol.* **4**:760-5) and inhibit angiogenesis (C.H. Yeh et al., 1998, *Blood* **92**:3268-3276). Without wishing to be bound by theory, it is possible that the disintegrin-like domain of the Gene 216 polypeptide inhibits angiogenesis in the respiratory system. Gene 216 variants that have partly functional or non-functional disintegrin activity may lack anti-angiogenesis function. These Gene 216 variants may give rise to angiogenesis and inflammation in the respiratory system, a phenotype of asthma.

#### **EXAMPLE 9: Identification of the Mouse Homolog for Gene 216**

The mouse ortholog of Gene 216 was identified by TBLASTN analysis of Gene 216 against mouse dbEST. BLAST analysis identified three mouse ESTs that were partially homologous to the human sequence but were not 100% homologous to any known mouse ADAM genes. The three mouse ESTs were 100% identical to a partially sequenced mouse BAC (BAC389B9; Accession Number AF155960). This BAC maps to mouse chromosome 2 in a region that is syntenic to human chromosome 20p13. The 47 kb BAC sequence was analyzed for potential genes using the Genscan gene prediction program (Burge and Karlin, *J. Mol. Biol.*, **268**:78-94). Additional putative exons were identified based on comparison of the human Gene 216 protein to the mouse BAC by TBLASTN. The results identified a mouse gene that contained an ORF of 2124 bp encoding a protein of 707 amino acids. The genomic nucleotide sequence of the mouse homolog is depicted in Figure 21 and the corresponding amino acid sequence is depicted in Figure 22. The mouse

amino acid sequence was analyzed by BLASTP analysis and found to have homology to mouse and human ADAM proteins. The mouse amino acid sequence was aligned against the amino acid sequence of human Gene 216 (BestFit, <http://www.gcg.com>) (Figure 23). The results showed that the mouse and human proteins shared ~70% identity at the amino acid level. This indicated that the mouse sequence was the murine ortholog of human Gene 216.

#### **EXAMPLE 10: Polymorphism Identification**

Polymorphisms were identified in the chromosome 20 region and subsequently used in association studies. Most of the data focused on the region of Gene 216.

1. Single Nucleotide Polymorphism (SNP) Discovery: An efficient tiered approach was used for mutation analysis. First, PCR assays were developed across exons to include the consensus splice sites. Assays were designed for all exons that contribute to the open reading frame of the gene. This strategy ensured the detection of mutations that would result in the modification of the protein sequence as well as mutations that would be predicted to disrupt mRNA splicing. The identified promoter and putative regulatory element for Gene 216 and a large intronic region were assayed for polymorphisms as well. Second, a total of 77 individuals were tested for polymorphisms using fluorescent SSCP (single strand conformational polymorphism). This sample size provided a 99% power to detect a polymorphism with a frequency of 3% or greater. Briefly, PCR was used to generate templates from asthmatic individuals that showed increased sharing for the 20p13-p12 chromosomal region and contributed towards linkage. Non-asthmatic individuals were used as controls. Enzymatic amplification of Gene 216 was accomplished using PCR with oligonucleotides flanking each exon as well as the putative 5' region. Primers were chosen to amplify each exon as well as 15 or more base pairs within each intron on either side of the splice site. The forward and the reverse primers were labeled with two different dye colors to allow analysis of each strand and confirm variants independently.



Standard PCR assays were utilized for each exon primer pair following optimization. Buffer and cycling conditions were specific to each primer set.

The products were denatured using a formamide dye and electrophoresed on non-denaturing acrylamide gels with varying concentrations of glycerol (at least  
5 two different glycerol concentrations).

Primers utilized in fluorescent SSCP experiments to screen coding and non-coding regions of Gene 216 for polymorphisms are provided in Table 8. Column 1 lists the genes targeted for mutation analysis. Column 2 lists the specific exons analyzed. Column 3 lists the primer names. Columns 4 and 5  
10 list the forward primer sequences and corresponding SEQ ID NOS, respectively. Columns 5 and 6 list the reverse primer sequences and corresponding SEQ ID NOS, respectively

**TABLE 8**

Gene	Exon	Assay Name	Primer Sequence	SEQ ID NO:	Primer Sequence	SEQ ID NO:
216	216_A	502_216_A_F_503_216_A_R	Ctgcctagagccgagga	63	agctctgagcagaacccatc	106
216	216_A	1623_216_A_F_1624_216_A_R	Caggagaccacggaagatcg	64	ctcaggggggtggagctg	107
216	216_A	1625_216_A_F_1626_216_A_R	Ttgcctgaacctctatcc	65	gagaggaggagagaaccgct	108
216	216_B	293_216_B_F_294_216_B_R	Cccctgtgttctcaggtc	66	agtgcacttggtgttctggg	109
216	216_C	295_216_C_F_296_216_C_R	Gctccacacttcttctgcc	67	tgatcctgcaccctctctg	110
216	216_D	297_216_D_F_298_216_D_R	Aggcaggaggaaagctaat	68	aagaggaggaggtgtgtagg	111
216	216_E	1290_216_E_F_1291_216_E_R	Cctaccacacctccctctt	69	gtgatcaggccactagggtg	112
216	216_F	299_216_F_F_300_216_F_R	Cctaccctctgcacctta	70	atacagcatccacctccca	113
216	216_G	301_216_G_F_302_216_G_R	aacttctcttgaggctgg	71	gaaggcagaatcccggt	114
216	216_H	700_216_H_F_701_216_H_R	cacacctgtgtgaggagaga	72	caccagcaacctgctgtc	115
216	216_I	305_216_I_F_306_216_I_R	ccacgaaggaccaccg	73	gggtcagaggcaccac	116
216	216_J	889_216_J_F_890_216_J_R	ctcactgggtgacctctg	74	gccgtagagcctctctgt	117
216	216_K	891_216_K_F_892_216_K_R	ctctacggccgcagtgac	75	gacgaccaagaacgcag	118
216	216_L	311_216_L_F_312_216_L_R	gtccctccatgcccaatg	76	tgagcggagagggaagt	119
216	216_L	313_216_L_F_314_216_L_R	cagggttaagtcggctgc	77	aaacctcaccctgaacctt	120
216	216_M	315_216_M_F_316_216_M_R	ctctctctgccttccccac	78	aagggtgctctgtctctct	121
216	216_N	317_216_N_F_318_216_N_R	tctactgtggggaagatggg	79	ccactcagctccactcccta	122
216	216_O	319_216_O_F_320_216_O_R	cccctctactctcccca	80	ggatcaaacggcaaggag	123
216	216_P	321_216_P_F_322_216_P_R	gaccttgggggttctaatcc	81	gctgagctctgagcaggtg	124
216	216_Q	323_216_Q_F_504_216_Q_R	gtgcacctgctcaggactc	82	gaaccgcaggagttagctc	125
216	216_R	325_216_R_F_326_216_R_R	cctggactcttatcacgtgc	83	atatgtctcagcaggagccc	126
216	216_S	327_216_S_F_328_216_S_R	ttaccctccaccattctcc	84	gcactctgtctccatgataa	127
216	216_S	1308_216_S_F_1309_216_S_R	gtggagagggaaggagaag	85	gaggcttgaatccaggtcc	128
216	216_T	1294_216_T_F_1295_216_T_R	ccccatgggtgaattaca	86	cagcaagacaccgcacttac	129
216	216_T	1296_216_T_F_1297_216_T_R	gcagctaggctacaggtaca	87	gggacagagggaaccattta	130
216	216_T	1298_216_T_F_1299_216_T_R	accacgcctatagccaacat	88	ttcttctgttcttccca	131

216	216_T	1300_216_T_F_1301_216_T_R	agggttagcactgggattgg	89	gtcctgggagtcgtgtgt	132
216	216_T	1302_216_T_F_1303_216_T_R	ccccaggaccactagcttct	90	aggaaaccagagccacacta	133
216	216_T	1304_216_T_F_1305_216_T_R	attgagctggagagtgtgcc	91	tgcctcgggtgagagtagc	134
216	216_T	1306_216_T_F_1307_216_T_R	ttcaagtctcctggagtgct	92	ttcctggatcactggctctc	135
216	216_AA	1619_216_AA_F_1620_216_AA_R	acaaggaccctctaaacgca	93	ttcagagcagtgaagaacct	136
216	216_PQ	1465_216_PQ_F_1466_216_PQ_R	acccttctgtgacaagccag	94	ctgggagtcggtagcaaca	137
216	216_QR	1467_216_QR_F_1468_216_QR_R	gtgtgtctaccgactcccag	95	aggccactggaacctcct	138
216	216_QR	1469_216_QR_F_1470_216_QR_R	cccagggtgcagagagcag	96	gcagcatggtacagggactg	139
216	216_QR	1471_216_QR_F_1472_216_QR_R	gtcctctgttccactctct	97	cagctgaccagtgtgtatgga	140
216	216_QR	1473_216_QR_F_1474_216_QR_R	gccacttctctgcacaaat	98	tgtcagacatggccacagag	141
216	216_QR	1475_216_QR_F_1476_216_QR_R	ttctctgtgacctgggtgt	99	agggtcctcttagctgccac	142
216	216_QR	1477_216_QR_F_1478_216_QR_R	atttgggccagagatggg	100	aggccttgcatttctgtg	143
216	216_QR	1479_216_QR_F_1480_216_QR_R	ggcagaggagcaaggtgg	101	caaagaaccttgatgtccg	144
216	216_QR	1481_216_QR_F_1482_216_QR_R	atggcttgggaatcatcaagg	102	ctcagctcccttctgtctc	145
216	216_QR	1483_216_QR_F_1484_216_QR_R	tagagagaggaggtgccagc	103	ctgtgtgggccattttg	146
216	216_RS	1485_216_RS_F_1486_216_RS_R	aaagatggcccacacagg	104	ggagaaatgggtgagggtaa	147
216	216_ST	1487_216_ST_F_1488_216_ST_R	agaactctcatgagccagc	105	aaagccacagcttctccct	148
216	216_ST	1489_216_ST_F_1490_216_ST_R	aggtttctgggctcaggta	149	caggatcttggcatctggac	153
216	216_UP	1463_216_UP_F_1464_216_UP_R	gtaggtgtgcagagcagg	150	ctggcttgcacagaagggt	154
216	216_U	1292_216_U_F_1293_216_U_R	tgtggacctagaatggtgagc	151	ctggagcacagtggcagtta	155
216	216_V	1736_216_V_F_1737_216_V_R	caaagtcacacaacaagcgg	152	tttggctgcctcagtttc	156

Once polymorphisms were identified, multiple individuals representative of each SSCP pattern and two genomic controls were sequenced for polymorphism validation and to identify SNPs. The variants detected in the initial set of asthmatic and normal individuals were subject to fluorescent sequencing (ABI) using a standard protocol described by the manufacturer (Perkin Elmer). In cases where SSCP did not identify polymorphisms in Gene 216, sequence information was obtained from 16 individuals that were identical by descent (IBD) in the region, and from 4 controls to ensure that potential polymorphisms were identified.

Primers utilized in DNA sequencing for purposes of confirming polymorphisms detected using fluorescent SSCP are provided in Table 9. Column 1 lists the specific exons sequenced. Column 2 lists the forward primer names, column 3 lists the forward primer sequences, and column 4 lists the corresponding SEQ ID NOS. Column 5 lists the reverse primer names, column 6 lists the reverse primer sequences, and column 7 lists the corresponding SEQ ID NOS.

**TABLE 9**

Exon	Forward	Forward Seq	SEQ ID NO:	Reverse Name	Reverse Seq	SEQ ID NO:
216 A	MDSeq 101 216 A F	cctctcaggagtagaggccc	157	MDSeq 101 216 A R	ccaagcacacitgagagtc	177
216 A	MDSeq 175 216 A F	agcgggttctctctctctc	158	MDSeq 175 216 A R	agccatgccctctgcttt	178
216 A	MDSeq 213 216 A F	cctctcaggagtagaggccc	159	MDSeq 213 216 A R	cagcccaagcacactga	179
216 A	MDSeq 334 216 A F	atgttactgaggccgaagg	160	MDSeq 334 216 A R	cccatagctgtgagctctc	180
216 B	MDSeq 296 216 B F	cccttccagcctctctctt	161	MDSeq 296 216 B R	aaagcttcaggaccacaaa	181
216 C	MDSeq 297 216 C F	caggactgcaaacatctga	162	MDSeq 297 216 C R	atcttggtccctgccatc	182
216 D	MDSeq 61 216 D F	tccttggtcttcccca	163	MDSeq 61 216 D R	gagggagctctttcccca	183
216 E	MDSeq 245 216 E F	aggcaggaggaagctgaat	164	MDSeq 245 216 E R	ggaccaccagggaagctg	184
216 F	MDSeq 57 216 F F	cctctgcccccttctgt	165	MDSeq 57 216 F R	aaecccagctcccaagag	185
216 G	MDSeq 336 216 G F	cctgaatgtccagagctctga	166	MDSeq 336 216 G R	ctgtcacctctggaaggaaac	186
216 H	MDSeq 155 216 H F	ggcctcagatcccaatttt	167	MDSeq 155 216 H R	actgcaggagggcccaag	187
216 I	MDSeq 363 216 I F	agagctctctgtctctct	168	MDSeq 363 216 I R	accgaacitgaaccacacc	188
216 J	MDSeq 181 216 J F	tcgccctcagcttctcag	169	MDSeq 181 216 J R	tggaggacgacccaagaagac	189
216 K	MDSeq 182 216 K F	tcacgtgggtgcctctga	170	MDSeq 182 216 K R	caagatcacacaacaagcgg	190
216 L	MDSeq 106 216 L F	gggttacttccccctcttg	171	MDSeq 106 216 L R	gaacctpagggccaccaatta	191
216 M	MDSeq 337 216 M F	ctgggcttccacccctgg	172	MDSeq 337 216 M R	ttggccttagttaattgtgc	192
216 N	MDSeq 338 216 N F	ctgggcttccacccctgg	173	MDSeq 338 216 N R	ttggccttagttaattgtgc	193
216 O	MDSeq 49 216 O F	tcaggttggtgaactctgc	174	MDSeq 49 216 O R	ctggagacnctggcagcta	194
216 P	MDSeq 248 216 P F	tagaatggtagcctctgcc	175	MDSeq 248 216 P R	aggagtaggctcagggaagca	195
216 Q	MDSeq 96 216 Q F	gaacttgggttctctaacc	176	MDSeq 96 216 Q R	tgtacttggaggttagaggcc	196
216 R	MDSeq 50 216 R F	agagggtgacttggagcaga	197	MDSeq 50 216 R R	ccagaaacctgattaggagg	219
216 S	MDSeq 262 216 S F	aggcaataaccactcagga	198	MDSeq 262 216 S R	tacctctcaccagaggcagg	220
216 T	MDSeq 255 216 T F	cccateggtgaatttata	199	MDSeq 255 216 T R	gccagaagctagtgtctctg	221
216 T	MDSeq 256 216 T F	gcctctggtgaccolctaac	200	MDSeq 256 216 T R	gcaggcagcttgggaagttt	222
216 T	MDSeq 257 216 T F	actcagtcgaaccataggcc	201	MDSeq 257 216 T R	ttatctggagacacagatgc	223
216 T	MDSeq 258 216 T F	tgtgtacccttgcctctgg	202	MDSeq 258 216 T R	gacctggatcaaaagctcc	224
216 T	MDSeq 358 216 T F	gcataagcaatgggagaat	203	MDSeq 358 216 T R	atgttggctatagagctgtgt	225
216 T	MDSeq 365 216 T F	actcagtcgaaccataggcc	204	MDSeq 365 216 T R	ttatctggagacacagatgc	226
216 U	MDSeq 244 216 U F	gcagggaaggtgtcatgtct	205	MDSeq 244 216 U R	ctgagtgaggaggacagaag	227
216 U	MDSeq 292 216 U F	gcagggaaggtgtcatgtct	206	MDSeq 292 216 U R	ctgagtgaggaggacagaag	228
216 V	MDSeq 389 216 V F	gggcatggagagggaag	207	MDSeq 389 216 V R	ccatagatcgccacag	229
216 AA	MDSeq 360 216 AA F	tcgtcctccagattcaagt	208	MDSeq 360 216 AA R	atttcaaggtcgaatgagg	230
216 PQ	MDSeq 300 216 PQ F	agaatgccttcaggagctt	209	MDSeq 300 216 PQ R	acttcttccatggcctctg	231
216 QR	MDSeq 301 216 QR F	gtgtgtctacccagctccag	210	MDSeq 301 216 QR R	accacccaggttcaagagaa	232
216 QR	MDSeq 303 216 QR F	ctgcttctgagcctackcc	211	MDSeq 303 216 QR R	tcccaagaccaggtatgtc	233
216 QR	MDSeq 321 216 QR F	aacaggaggttcagtgcc	212	MDSeq 321 216 QR R	ctggggatgagaagcagc	234
216 QR	MDSeq 322 216 QR F	agcgaggtgtgattgaggt	213	MDSeq 322 216 QR R	cttctcccttccctctcac	235
216 QR	MDSeq 361 216 QR F	tgttcaggctgaagtatgc	214	MDSeq 361 216 QR R	attttgcagagggaagtgcc	236
216 QR	MDSeq 362 216 QR F	gccacttctctgcacaaat	215	MDSeq 362 216 QR R	catttctccaggtctgac	237
216 RS	MDSeq 339 216 RS F	ctgagcccagaaactgaat	216	MDSeq 339 216 RS R	tcagagcctggaggaaatgt	238
216 ST	MDSeq 302 216 ST F	gtgagtgaggcaccagg	217	MDSeq 302 216 ST R	gttctggagtggtgtggt	239
216 UP	MDSeq 359 216 UP F	cctagatggcagggaagtga	218	MDSeq 359 216 UP R	caggagctggttagcaaca	240

Single nucleotide polymorphisms (SNPs) that were identified in Gene

216 are provided in Table 10. Column 1 lists the SNP numbers (1-48).

Column 2 lists the exons that either contain the SNPs or are flanked by intronic sequences that contain the SNPs. Column 3 lists the PMP sites for the SNPs.

A "-" denotes polymorphisms which are 5' of the exon that are within the intronic region. The corresponding number is given from the 3' to 5' direction.

A "+" denotes polymorphisms which are 3' of the exon that are within the intronic region. The number corresponding to the "+" is given from the 5' to 3' direction. Columns 2 and 3, combined, show the SNP names as described herein, e.g., T+1, T+2, etc. Column 4 indicates whether the SNP was detected

in an exon or intron sequence. Column 5 lists the SNP locations in the Gene 216 genomic sequence of SEQ ID NO:6 (Figure 7). Column 6 lists the SNP reference sequences which illustrate the SNP nucleotide changes with underlining. Column 7 lists the SEQ ID NOs of the SNP reference sequences.

5 Column 8 lists the base changes of the SNP sequences. Column 9 lists the amino acid changes resulting from the SNP sequences.

**TABLE 10**

SNP	Exon	PMP site	Location	Location	Sequence (20nt+SNP+20nt)	SEQ ID NO:	PMP	AA Change
1	A	-1	Intron	4653	GCCCTCTGAGACCGACGGGGAGGGACGGCTCGGGCCGGTC	241	A>T	
2	A	-2	Intron	4610	CAAGAACCTTCCCAGCGGTCTCTCTCTCTCAGGAGTAG	242	C>A	
3	C	-1	Intron	9827	CACCATCTCAGCTCCACACTCTTTCTTGCCCAGGTCTCGAA	243	C>T	
4	C	-2	Intron	9826	CCACCATCTCAGCTCCACACTCTTTCTTGCCCAGGTCTCGA	244	T>A	
5	D	-1	Intron	11687	ACAACCTAAGCCATCAACAAGCTCTCTCTAGCCCCAAG	245	G>C	
6	D	-2	Intron	11661	TGGTGCTTCCCATATTCACATCTCCCACTAAGCCATCA	246	T>C	
7	D	1		11912	CAGGATACATAGAAACCCACTACGGCCCGAGATGGGCAGCC	247	T>C	Tyr>His
8	F	+1	Intron	12545	CCCTCCAAATCAGAAGAGACAGGAATTCACAGGCCTCGAG	248	A>G	
9	F	1		12411	AGCTGCTCACCTGGAAAGGAACCTGTGGCCACAGGGATCC	249	A>G	Thr>Ala
10	G	-1	Intron	12637	ACTTCTCTCTGGGAGCTGGGGTTGGGGGTCAAGGCTCAAGC	250	G>A	
11	I	1		13197	TTCTGTCAGTGGCGCGGGGGCTGTGGGGCGCAGCGGCCCC	251	G>A	none
12	L	+1	Intron	14481	GGTTCAAGGTGAGGGTTTCGGGGAGCTTGGGAGCCGGCCT	252	G>T	
13	L	-1	Intron	14043	CAGAGAAGCGCGGGGGTTGGGGGACTGTCCCTCCATGCC	253	G>A	
14	L	-2	Intron	13988	CCCCCTCTGGGCTCTGCGGCTGTGGCGGCTGTAGCCAAGC	254	G>A	
15	L	1		14135	CAGCCGCGCCAGCTGCGCGCTCTTCCGCAAGGGGGGC	255	C>T	Ala>Val
16	Q	+1	Intron	16158	AGTGGCTCCAGTCAAGCGAGGGGGTGGATCCCTGCCCC	256	A>T	
17	Q	1		15865	TGCTGGCCATGCTCCTCAGCGTCTGCTGCTCTGCTCCCA	257	G>A	Val>Ile
18	Q	2		15888	CTGCTGCTCTGCTCCAGGGGCCGGCTGGCTGGTGTG	258	G>C	
19	QR	+1	Intron	16133	GAAGTAGCTTTGAACAGGAGGTTCCAGTGGCCTCCAGTCA	259	G>T	
20	QR	+3	Intron	16361	GCCTCTGTCTACCAAGTTTGGGCCCTTGCCACTTCTCT	260	C>T	
21	QR	+4	Intron	16404	ACAAATCACCTCTGTACCCCTTGAAGTTCCCAAATGCTG	261	C>A	
22	QR	+5	Intron	16465	TCCATACCACTGGTCAGCTGCGGTGCTGGCTGCCCTGTGC	262	C>T	
23	QR	+6	Intron	16486	GGTGCTGGCTGCCCTGTGCCAGGGCCCTGCCTTAACCCAG	263	C>T	
24	QR	+7	Intron	16936	GGAAATGACAAGGCCTTGGGGGATGGGATGGGGACAGTCA	264	G>A	
25	R	+1	Intron	17510	AGGGCTCATGCCTCTGCTCTCTCCAGATGGGCAGCACCC	265	C>T	
26	R	+2	Intron	17571	GCCCCCCCCAGCCCCAGGGTCTCTGCTGACCATATTAC	266	T>G	
27	R	1		17403	CCTGGGCGGCGTTCAACCCATGGAGTTGGGCCCCACAGCC	267	T>C	Met>Thr
28	R	2		17432	GCCCCACAGCCACTGGACAGCCCTGGCCCTGGGTGAGTG	268	C>T	Pro>Ser
29	RS	-1	Intron	17451	GCCCTGGCCCCCTGGGTGAGTGGAGCACCAGGGGGAGGTGG	269	G>T	

30	T	+1	intron	17958	TGCAGCCTGGGGCCCCAGTCCTTAGGGGACAACATATCCTC	270	C>A	
31	T	+2	intron	17924	CACTGAGTGAGGATGGGCTCTCTGCCACACAGCTTGCAGCC	271	T>C	
32	T	+3	intron	17916	CTGGTCCTCACTGAGTGAGGATGGGCTCTCTGCCACACAGC	272	A>G	
33	T	+4	intron	17834	ATGACCTCTTGGTTATCATGGAGACCAGGATGCTGGAAGCC	273	G>C	
34	T	1	3' UTR	18833	AGCAAGACACCCGATCTACA <del>G</del> AAAAATTTAAAAATTAGCTG	274	G>A	
35	T	2		18787	GGAGGATCACCAGAGGCCAG <del>C</del> AGGTCCACACCAGCCTGGG	275	C>G	
36	T	3	3' UTR	18760	ATCCAGCACTTTGGGAAGC <del>C</del> GGGGTAGGAGGATCACCAG	276	C>T	
37	T	4	3' UTR	18497	AGCCTGGCTGGCCTCTGCAA <del>A</del> CAAACATAATTTGGGGACC	277	A>G	
38	T	5		18476	ACTGAGTCCACACTCCCCTG <del>C</del> AGCCTGGCTGGCCTCTGCAA	278	C>G	
39	T	6		18206	TCCAGGAACCCAGAGCCACA <del>T</del> TAGAAGTTCCTGAGGGCTG	279	T>C	
40	T	7		18174	TTCTTCCCCGAGTGAGGCTT <del>C</del> GACCCACCCACTCCAGGAAC	280	C>T	
41	T	8		17997	TCCTCATTCTCAGCAGATCA <del>A</del> GTCCAGATGCCAAGATCCTG	281	A>T	Gln>His
42	T	-2	intron	19094	CTGAGGACCACACGGGGTGG <del>T</del> GGTTGGCGGGGTGGTGGTT	282	T>C	
43	T	-4	intron	19160	GGCTGGCAGGCCGAGCCTAG <del>A</del> TGGCAGCCAGAGCCCCAGG	283	A>G	
44	T	-5	intron	19244	CTTTGCTCTGTCACTCCTGC <del>C</del> TCCTTGGGCGTTCACATTC	284	C>T	
45	U	-1	intron	15423	GTGAGCTCTGCCACCCGAC <del>C</del> CCTCCTTGCGTTTGAATCC	285	C>T	
46	V	+1	intron	13859	TGGCGAGGTTACTCCTACAC <del>C</del> GGGAGGAGCACCGTCGGGT	286	C>T	
47	V	+2	intron	13921	GGCTGCTCACTATTGGGGCC <del>G</del> CATCGTCCCCTGTCCCGCTT	287	G>T	
48	V	+3	intron	13938	GCCGCATCGTCCCCTGTCCC <del>G</del> CTTGTGTGTGACTTTGCGC	288	G>A	

Using an in-house program called snp\_view; the genomic structure of the gene is diagrammatically shown in Figure 11. The exons are shown to scale and the SNPs are identified by their location along the genomic BAC DNA. The polymorphic sites identified in the Gene 216 genomic sequence are also shown by the underlined nucleotides in Figure 29. The polymorphic sites discovered within the cDNA and the corresponding amino acid position in Gene 216 are underlined in Figure 24. It will be understood by those of skill in the art that the SNPs identified in the Gene 216 genomic sequence can be correlated to the SNP positions identified in the Gene 216 cDNA sequence by aligning the genomic and cDNA sequences.

#### **EXAMPLE 11: Polymorphism Genotyping**

Once putative variants were confirmed by sequencing, rapid allele specific assays were designed to type more than 400 individuals (> 200 cases and > 200 controls) for use in the association studies. All coding SNPs (cSNPs) that resulted in an amino acid change were typed. Neutral

polymorphisms were typed if: 1) the polymorphism was present in an exon lacking a cSNP that resulted in an amino acid change; 2) the polymorphism was present in an exon containing a cSNP resulting in an amino acid change but the two polymorphisms were observed to have different frequencies; and  
5 3) the polymorphism was in an intronic region adjacent to an exon without a cSNP. If results from the association studies appeared positive, additional neutral polymorphisms were typed. More than 30 allele specific assays from Gene 216 were typed for the case control population (Table 11).

Two types of allele specific assays (ASAs) were used. If the SNP  
10 resulted in a mutation that created or abolished a restriction site, restriction fragment length polymorphisms (RFLPs) were obtained from PCR products that spanned the variants, and the RFLPs were analyzed. If the polymorphisms did not result in RFLPs, allele specific oligonucleotide assays were used. For these assays, PCR products that spanned the polymorphism  
15 were electrophoresed on agarose gels and transferred to nylon membranes by Southern blotting. Oligomers 16-20 bp in length were designed such that the middle base was specific for each variant. The oligomers were labeled and successively hybridized to the membrane in order to determine genotypes. The specific method used to type each SNP is indicated in Table 11.

20 Table 11 below contains the information relating to the specific assay used. Column 1 lists the SNP designation number. Column 2 lists the specific assay used, either RFLP or ASO. Column 3 lists the enzyme used in the RFLP assay (described below). Columns 4 and 6 list the sequence of the primers used in the ASO assay (described below). Columns 5 and 7 list the  
25 corresponding SEQ ID NOS for the primers.

1. RFLP Assay: The amplicon containing the polymorphism was PCR amplified using primers that were used to generate a fragment for sequencing (sequencing primers) or SSCP (SSCP primers). The appropriate population of individuals was PCR amplified in 96 well microtitre plates.

30 Enzymes were purchased from NEB. The restriction cocktail containing the appropriate enzyme for the particular polymorphism is added to the PCR

product. The reaction was incubated at the appropriate temperature according to the manufacturer's recommendations (NEB) for 2-3 hr, followed by a 4°C incubation. After digestion, the reactions were size fractionated using the appropriate agarose gel depending on the assay specifications (2.5%, 3%, or Metaphor, FMC Bioproducts). Gels were electrophoresed in 1 X TBE Buffer at 170 Volts for approximately 2 hr. The gel was illuminated using ultraviolet light and the image was saved as a Kodak 1D file. Using the Kodak 1D image analysis software, the images were scored and the data was exported to Microsoft EXCEL (<http://www.microsoft.com>).

- 10           2.     ASO assay: The amplicon containing the polymorphism was PCR amplified using primers that were used to generate a fragment for sequencing (sequencing primers) or SSCP (SSCP primers). The appropriate population of individuals was PCR amplified in 96 well microtitre plates and re-arrayed into 384 well microtitre plates using a Tecan Genesis RSP200. The amplified products were loaded onto 2% agarose gels and size fractionated at 150V for 5 min. The DNA was transferred from the gel to Hybond N+ nylon membrane (Amersham-Pharmacia) using a Vacuum blotter (Bio-Rad). The filter containing the blotted PCR products was transferred to a dish containing 300 ml pre-hybridization solution (5 X SSPE (pH 7.4), 2% SDS, 5 X Denhardt's). The filter was incubated in pre-hybridization solution at 40°C for over 1 hr. After pre-hybridization, 10 ml of the pre-hybridization solution and the filter were transferred to a washed glass bottle. The allele specific oligonucleotides (ASO) were designed with the polymorphism in the middle. The size of the oligonucleotide was dependent upon the GC content of the sequence around the polymorphism. Those ASOs that had a G or C polymorphism were designed so that the  $T_m$  was between 54-56°C and those that had an A or T variance were designed so that the  $T_m$  was between 60-64°C. All oligonucleotides were phosphate free at the 5' end and purchased from GibcoBRL. For each polymorphism, 2 ASOs were designed: one for each variant.
- 25  
30

The two ASOs that represented the polymorphism were resuspended

at a concentration of 1  $\mu\text{g}/\mu\text{l}$  and separately end-labeled with  $\gamma\text{-ATP}^{32}$  (6000 Ci/mmol) (NEN) using T4 polynucleotide kinase according to manufacturer recommendations (NEB). The end-labeled products were removed from the unincorporated  $\gamma\text{-ATP}^{32}$  by passing the reactions through Sephadex G-25  
5 columns according to manufacturers recommendation (Amersham-Pharmacia). The entire end-labeled product of one ASO was added to the bottle containing the appropriate filter and 10 ml hybridization solution. The hybridization reaction was placed in a rotisserie oven (Hybaid) and left at 40°C for a minimum of 4 hr. The other ASO was stored at -20°C.

10 After the prerequisite hybridization time had elapsed, the filter was removed from the bottle and transferred to 1 L of wash solution (0.1 X SSPE (pH 7.4), 0.1% SDS) pre-warmed to 45°C. After 15 min, the filter was transferred to another L of wash solution (0.1 X SSPE (pH 7.4), 0.1% SDS) pre-warmed to 50°C. After 15 min, the filter was wrapped in Saran, placed in  
15 an autoradiograph cassette and an X-ray film (Kodak) placed on top of the filter. Typically, an image would be observed on the film within 1 hr. After an image had been captured on film for the 50°C wash, the process was repeated for wash steps at 55°C, 60°C and 65°C. The image that captured the best result was used.

20 The ASO was removed from the filter by adding 1 L of boiling strip solution (0.1 x SSPE (pH 7.4), 0.1% SDS). This was repeated two more times. After removing the ASO the filter was pre-hybridized in 300 ml pre-hybridization solution (5 X SSPE (pH 7.4), 2% SDS, 5 X Denhardt's) at 40°C for over 1 hr. The second end-labeled ASO corresponding to the other variant  
25 was removed from storage at -20°C and thawed at room temperature. The filter was placed into a glass bottle along with 10 ml hybridization solution and the entire end-labeled product of the second ASO. The hybridization reaction was placed in a rotisserie oven (Hybaid, <http://www.hybaid.co.uk>) and left at 40°C for a minimum of 4 hr. After the hybridization, the filter was washed at  
30 various temperatures and images captured on film as described above.

The two films that best captured the allele-specific assay with the two



ASOs were converted into digital images by scanning them into Adobe PhotoShop. These images were overlaid against each other in Graphic Converter and then scored.

**TABLE 11**

SNP	ASA Type	RFLP Enzyme	ASO Primer1	SEQ ID NO:	ASO Primer2	SEQ ID NO:
1	ASO		gccgtcccaccccgctcg	289	gccgtcccaccccgctcg	299
2	ASO		cctcctctcttggcgac	290	tcctcctctattggcgacce	300
3	ASO		tccacactctttttgcc	291	ctccacacttttttggcca	301
4	ASO		gtccacactctttttgcc	292	gtccacactctttttgc	302
5	ASO		tcaccaaggctcctctct	293	tcaccaagcctcctctct	303
6	Alt. Meth					
7	RFLP	XcmI				
8	ASO		cagaagagacaggaattcaca	294	agaagagacgggaattcac	304
9	ASO		tggaaaggaaacctgtggcc	295	tggaaaggagcctgtgg	305
10	ASO					
11	ASO					
12	ASO		gggttcggggagcttg	296	agggtttcgtggagcttgg	306
13	ASO		gggttggggactgtc	297	gggttggggactgtcc	307
14	ASO		ctctgcgcgtctggcg	298	gctctgcgcctctggcg	308
15	RFLP	BssHII				
16	ASO		agtcaagcgagggggtgg	309	agtcaagcgagggggtgg	322
17	ASO		cctcagcgtcctgtg	310	ctcctcagcctcctgtg	323
18	RFLP	KasI				
19	ASO		aacaggagggttcagtg	311	gaacaggagtttcagtg	324
20	ASO		accagttttcggccctt	312	caccagtttttggcccttg	325
21	ASO		ctgtcaccctctgaagt	313	ctgtcaccctctgaagtc	326
22	ASO		tcagctcggtgctgg	314	ggtcagctggtgctgg	327
23	RFLP	BstNI				
24	ASO		gccttggggatgga	315	aggccttggggatggat	328
25	ASO		tectgectcttcag	316	tcctgectcttcag	329
26	RFLP	BglI				
27	RFLP	NcoI				
28	ASO		actggacagccctggc	317	actggacagtctggc	330
29	ASO					
30	RFLP	Bsu36I				
31	ASO		ctgtgtggcagagagccca	318	tggtggcaggagagccca	331
32	ASO					
33	RFLP	BsaI				
34	Alt. Meth					
35	RFLP	Cac8I				
36	RFLP	MspI				
37	ASO		aattatgtttgttcagaggc	319	attatgtttgttcagagg	332
38	RFLP	Fnu4HI				
39	ASO		gaacttctagtgtggctct	320	ggaacttctaagtgtggctctg	333
40	RFLP	TaqI				
41	RFLP	NlaIII				
42	ASO					
43	RFLP	StyI				

44	ASO		ccaagggaggcaggagt	321	ccaaggggaagcaggagtga	334
45	RFLP	HinfI				
46	RFLP	BsrI				
47	RFLP	Eco109 I				
48	ASO					

### **EXAMPLE 12: Association Study Analysis**

1. Case-Control Study: In order to determine whether polymorphisms in candidate genes were associated with the asthma phenotype, association studies were performed using a case-control study design. In a well-matched design, the case-control approach is more powerful than the family based transmission disequilibrium test (TDT) (N.E. Morton and A. Collins, 1998, *Proc. Natl. Acad. Sci. USA* **95**:11389-93). Case-control studies are, however, sensitive to population heterogeneity.

To avoid issues of population admixture, which can bias case-control studies, the unaffected controls were collected in both the US and the UK. A total of three hundred controls were collected, 200 in the UK and 100 in the US. Inclusion into the study required that the control individual was negative for asthma, as determined by self-report of never having asthma, had no first degree relatives with asthma, and was negative for eczema and symptoms indicative of atopy within the past 12 months. Data from an abbreviated questionnaire similar to that administered to the affected sib pair families were collected. Results from skin prick tests to 4 common allergens were also collected. The results of the skin prick test were used to select a subset of controls that were most likely to be asthma and atopy negative.

A subset of unrelated cases was selected from the affected sib pair families based on the evidence for linkage at the chromosomal location near a given gene. One affected sib demonstrating identity-by-descent (IBD) at the appropriate marker loci was selected from each family. Since the appropriate cases may vary for each gene in the chromosome 20 region, a larger collection of individuals who were IBD across a larger interval were genotyped, and a subset was used in the analyses. On average, 130 IBD affected individuals and 200 controls were compared for allele and genotype frequencies. This

number provided an 80% power to detect a difference of 5% or greater between the two groups for a rare allele ( $\leq 5\%$ ) at a 0.05 level of significance.

For a common allele (50%), the number provided an 80% power to detect a difference of 10% or more between the two groups.

5           For each polymorphism, the frequency of the alleles in the control and case populations was compared using a Fisher exact test. A mutation that increased susceptibility to the disease would be more prevalent in the cases than in the controls, while a protective mutation would be more prevalent in the control group. Similarly, the genotype frequencies of the SNPs were compared  
10       between cases and controls. P-values for both the allele and genotype were plotted against a coordinate system based on genomic sequence to visualize regions where allelic association was present. A small p-value (or a large value of  $-\log(p)$  as plotted in the figures described below) was indicative of an association between the SNPs and the disease phenotype. The analysis was  
15       repeated for the US and UK population separately to adjust for the possibility of genetic heterogeneity.

2.       Association test with individual SNPs: Chromosomal regions harboring asthma susceptibility genes were identified by association studies using the SNP typing data. Two separate phenotypes were used in these  
20       analyses: asthma and bronchial hyper-responsiveness.

a.       Asthma Phenotype: The significance levels (p-values) for allelic association of all typed SNPs in Gene 216 to the asthma phenotype are plotted in Figure 25 (combined population) and Figure 26 (US and UK populations separately). The most significant result in the combined population  
25       was observed for Gene 216 exon T+1, where 92.4% of the cases harbored the intronic mutation, while the SNP was present in only 85.2% of the controls ( $p = 0.0055$ ). Six additional SNPs in Gene 216 (T5, QR+7, QR+4, Q2, Q1, and U-1) were significant at the 0.05 level. Frequencies and p-values for SNPs associated with the asthma phenotype in Gene 216 are presented in Tables  
30       12, 13, and 14 for the combined population and for the UK and US populations, separately.

**TABLE 12**

Asthma Yes/NO						
Combined US and UK						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	66.5%	215	71.5%	128	0.2029	0.1482
gene216_T_3	8.7%	213	9.5%	131	0.7841	0.6895
gene216_T_4	96.3%	215	98.5%	129	0.1576	0.1513
gene216_T_5	76.7%	217	83.3%	129	<b>0.0420</b>	<b>0.0468</b>
gene216_T_6	77.8%	214	78.4%	125	0.9235	0.9791
gene216_T_7	96.3%	215	98.5%	129	0.1576	0.1513
gene216_T_8	96.5%	211	98.1%	129	0.2528	0.2456
gene216_T_+1	85.2%	216	92.4%	131	<b>0.0055</b>	<b>0.0178</b>
gene216_T_+2	37.3%	209	39.0%	127	0.6825	0.7722
gene216_T_+4	24.4%	215	26.3%	131	0.5886	0.7410
gene216_R_+2	88.3%	217	88.9%	131	0.8076	0.9005
gene216_R_+1	88.7%	191	88.8%	120	1.0000	0.8394
gene216_R_2	9.4%	208	10.8%	125	0.5928	0.7656
gene216_R_1	11.3%	217	11.8%	131	0.9025	0.7483
gene216_QR_+7	78.1%	215	85.7%	129	<b>0.0160</b>	<b>0.0265</b>
gene216_QR_+6	0.5%	216	0.8%	129	0.6323	0.6317
gene216_QR_+5	46.4%	210	48.8%	129	0.5794	0.4165
gene216_QR_+4	51.5%	205	59.9%	126	<b>0.0367</b>	0.1272
gene216_Q_+1	51.2%	206	52.5%	120	0.8075	0.6608
gene216_Q_2	73.7%	217	80.5%	131	<b>0.0432</b>	0.0831
gene216_Q_1	89.5%	209	94.8%	125	<b>0.0213</b>	0.0584
gene216_U_-1	85.0%	217	91.2%	131	<b>0.0184</b>	0.0659
gene216_L_+1	88.7%	213	88.9%	131	1.0000	0.9672
gene216_L_1	99.3%	217	99.6%	131	1.0000	1.0000
gene216_L_-1	88.9%	212	89.2%	130	1.0000	1.0000
gene216_L_-2	92.9%	212	93.1%	131	1.0000	0.9379
gene216_V_+2	71.3%	216	77.1%	129	0.1085	0.2262
gene216_V_+1	96.1%	217	97.2%	125	0.5223	0.5145
gene216_I_1	84.9%	212	85.3%	129	0.9124	1.0000
gene216_G_-1	90.7%	210	91.3%	127	0.8900	0.7683
gene216_F_+1	65.2%	197	70.4%	120	0.1913	0.4109
gene216_F_1	96.8%	217	96.9%	129	1.0000	1.0000
gene216_D_1	0.0%	215	0.4%	131	0.3786	0.3786
gene216_D_-2	0.7%	214	0.8%	127	1.0000	1.0000

**TABLE 13**

Asthma Yes/No						
UK population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	65.8%	139	74.3%	101	0.0566	0.1266
gene216_T_3	8.3%	139	9.6%	104	0.6308	0.7329
gene216_T_4	97.1%	140	98.5%	103	0.3689	0.3633
gene216_T_5	75.4%	140	83.3%	102	<b>0.0426</b>	<b>0.0365</b>
gene216_T_6	78.5%	137	80.1%	98	0.7301	0.8875
gene216_T_7	97.5%	138	99.0%	102	0.3129	0.3082
gene216_T_8	97.8%	137	98.5%	102	0.7388	0.7363
gene216_T_+1	86.4%	140	93.8%	104	<b>0.0105</b>	<b>0.0243</b>
gene216_T_+2	37.9%	136	40.5%	100	0.5682	0.8375
gene216_T_+4	25.2%	139	26.0%	104	0.9163	0.6037
gene216_R_+2	87.5%	140	87.5%	104	1.0000	1.0000
gene216_R_+1	86.9%	122	91.1%	95	0.2211	0.4281
gene216_R_2	10.5%	134	8.2%	98	0.4279	0.7007
gene216_R_1	13.2%	140	8.7%	104	0.1473	0.3472
gene216_QR_+7	79.5%	139	86.4%	103	0.0535	0.1362
gene216_QR_+6	0.0%	139	1.0%	103	0.1806	0.1801
gene216_QR_+5	44.4%	133	50.0%	102	0.2273	0.2470
gene216_QR_+4	48.1%	128	59.1%	99	<b>0.0229</b>	0.0730
gene216_Q_+1	53.1%	129	50.5%	97	0.6346	0.5458
gene216_Q_2	72.9%	140	84.6%	104	<b>0.0020</b>	<b>0.0050</b>
gene216_Q_1	89.4%	132	95.1%	101	<b>0.0274</b>	0.0732
gene216_U_-1	86.1%	140	92.3%	104	<b>0.0419</b>	0.0763
gene216_L_+1	87.0%	138	91.8%	104	0.1059	0.2969
gene216_L_1	99.3%	140	99.5%	104	1.0000	1.0000
gene216_L_-1	87.2%	137	92.2%	103	0.0992	0.1655
gene216_L_-2	92.7%	137	92.3%	104	0.8633	1.0000
gene216_V_+2	71.6%	139	79.1%	103	0.0717	0.1519
gene216_V_+1	97.1%	140	98.0%	99	0.7685	0.7655
gene216_I_1	83.7%	138	89.2%	102	0.1094	0.1323
gene216_G_-1	90.2%	137	90.1%	101	1.0000	0.4913
gene216_F_+1	64.1%	128	74.2%	93	<b>0.0295</b>	0.0711
gene216_F_1	97.9%	140	98.0%	102	1.0000	1.0000
gene216_D_1	0.0%	139	0.5%	104	0.4280	0.4280
gene216_D_-2	0.7%	139	1.0%	101	1.0000	1.0000

**TABLE 14**

Asthma Yes/No						
US population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	67.8%	76	61.1%	27	0.4053	0.1776
gene216_T_3	9.5%	74	9.3%	27	1.0000	1.0000
gene216_T_4	94.7%	75	98.1%	26	0.4519	0.4404
gene216_T_5	79.2%	77	83.3%	27	0.5583	0.7765
gene216_T_6	76.6%	77	72.2%	27	0.5819	0.6932
gene216_T_7	94.2%	77	96.3%	27	0.7320	0.7241
gene216_T_8	93.9%	74	96.3%	27	0.7308	0.7226
gene216_T_+1	82.9%	76	87.0%	27	0.5262	0.8281
gene216_T_+2	36.3%	73	33.3%	27	0.7416	0.5739
gene216_T_+4	23.0%	76	27.8%	27	0.5795	0.6743
gene216_R_+2	89.6%	77	94.4%	27	0.4127	0.3874
gene216_R_+1	92.0%	69	80.0%	25	<b>0.0334</b>	<b>0.0361</b>
gene216_R_2	7.4%	74	20.4%	27	<b>0.0188</b>	<b>0.0208</b>
gene216_R_1	7.8%	77	24.1%	27	<b>0.0030</b>	<b>0.0055</b>
gene216_QR_+7	75.7%	76	82.7%	26	0.3410	0.0921
gene216_QR_+6	1.3%	77	0.0%	26	1.0000	1.0000
gene216_QR_+5	50.0%	77	44.4%	27	0.5287	0.6337
gene216_QR_+4	57.1%	77	63.0%	27	0.5218	0.4709
gene216_Q_+1	48.1%	77	60.9%	23	0.1345	0.3169
gene216_Q_2	75.3%	77	64.8%	27	0.1571	0.1404
gene216_Q_1	89.6%	77	93.8%	24	0.5726	1.0000
gene216_U_-1	83.1%	77	87.0%	27	0.6654	0.8280
gene216_L_+1	92.0%	75	77.8%	27	<b>0.0116</b>	<b>0.0123</b>
gene216_L_1	99.4%	77	100.0%	27	1.0000	1.0000
gene216_L_-1	92.0%	75	77.8%	27	<b>0.0116</b>	<b>0.0123</b>
gene216_L_-2	93.3%	75	96.3%	27	0.7362	0.5089
gene216_V_+2	70.8%	77	69.2%	26	0.8614	0.8889
gene216_V_+1	94.2%	77	94.2%	26	1.0000	1.0000
gene216_I_1	87.2%	74	70.4%	27	<b>0.0105</b>	<b>0.0074</b>
gene216_G_-1	91.8%	73	96.2%	26	0.3635	0.3440
gene216_F_+1	67.4%	69	57.4%	27	0.2401	0.3270
gene216_F_1	94.8%	77	92.6%	27	0.5136	0.5043
gene216_D_1	0.0%	76	0.0%	27	1.0000	1.0000
gene216_D_-2	0.7%	75	0.0%	26	1.0000	1.0000

- b. Bronchial Hyper-responsiveness: The analyses were repeated using asthmatic children with borderline to severe BHR ( $PC_{20} \leq 16$  mg/ml) or  $PC_{20}(16)$ , as described in the linkage section. First, sibling pairs were identified where both sibs were affected and satisfied this new criteria.
- 5 Of these pairs, one sib was included in the case/control analyses if they showed evidence of linkage at the gene of interest. This phenotype was more restrictive than the Asthma yes/no criteria; hence the number of cases included in the analyses was reduced approximately in half. If the  $PC_{20}(16)$  subgroup represented a more genetically homogeneous sample, one expected to see an
- 10 increase in the effect size compared to the one observed in the original set of cases. However, the reduction in sample size could result in estimates that were less accurate and that could obscure a trend in allele frequencies in the control group, the original set of cases and the  $PC_{20}(16)$  subgroup. In addition, the reduction in sample size could induce a reduction in power (and increase
- 15 in p values) in spite of the larger effect size.

The significance levels (p-values) for allelic association of all typed SNPs in Gene 216 to the BHR phenotype are plotted in Figure 27 (combined population) and Figure 28 (US and UK populations separately). Frequencies and p-values for SNPs associated with the BHR phenotype in Gene 216 are

20 presented in Tables 15, 16, and 17 for the combined population and for the UK and US populations, separately. Again, multiple SNPs in Gene 216 were associated with the phenotype in each separate population. In the UK population, the most significant SNP was in Gene 216, exon Q2, where 87% of the cases had the mutation compared to 72.9% for the controls ( $p = 0.0038$ ).

25 For the US population, the most significant association was found with the SNP in Gene 216 exon R 1, where 28.6% of the cases carried the mutation compared to 7.8% for the controls ( $p = 0.0041$ ).

In summary, Gene 216 associated with the phenotypes of both asthma and bronchial hyper-responsiveness. Association was found with multiple

30 SNPs in both the UK and US populations. The 3' region of the gene, which contains the transmembrane domain, the cytoplasmic domain, and the 3' UTR,

appeared to have the strongest association. Taken together, these data strongly suggested that Gene 216 is an asthma susceptibility gene.



**TABLE 15**

BHR						
Combined US and UK						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	66.5%	215	67.7%	62	0.8294	0.1358
gene216_T_3	8.7%	213	9.4%	64	0.8592	0.6092
gene216_T_4	96.3%	215	98.4%	62	0.3878	0.3797
gene216_T_5	76.7%	217	79.8%	62	0.5428	0.5315
gene216_T_6	77.8%	214	78.3%	60	1.0000	0.8426
gene216_T_7	96.3%	215	97.7%	64	0.5856	0.5786
gene216_T_8	96.5%	211	97.6%	63	0.7758	0.7721
gene216_T_+1	85.2%	216	90.6%	64	0.1413	0.3117
gene216_T_+2	37.3%	209	41.8%	61	0.3978	0.6939
gene216_T_+4	24.4%	215	26.6%	64	0.6421	0.2498
gene216_R_+2	88.3%	217	88.3%	64	1.0000	0.8975
gene216_R_+1	88.7%	191	89.2%	60	1.0000	0.7540
gene216_R_2	90.6%	208	91.1%	62	1.0000	1.0000
gene216_R_1	11.3%	217	11.7%	64	0.8750	0.7576
gene216_QR_+7	78.1%	215	82.0%	64	0.3876	0.1711
gene216_QR_+6	99.5%	216	100.0%	63	1.0000	1.0000
gene216_QR_+5	46.4%	210	46.8%	63	1.0000	0.5530
gene216_QR_+4	51.5%	205	58.9%	62	0.1521	0.3393
gene216_Q_+1	51.2%	206	51.8%	57	1.0000	0.7632
gene216_Q_2	73.7%	217	79.7%	64	0.2009	0.0664
gene216_Q_1	89.5%	209	94.2%	60	0.1565	0.4299
gene216_U_-1	85.0%	217	89.8%	64	0.1915	0.5304
gene216_L_+1	88.7%	213	89.8%	64	0.8722	0.9410
gene216_L_1	0.7%	217	0.8%	64	1.0000	1.0000
gene216_L_-1	88.9%	212	89.1%	64	1.0000	1.0000
gene216_L_-2	7.1%	212	8.6%	64	0.5661	0.5313
gene216_V_+2	71.3%	216	75.0%	64	0.4343	0.7291
gene216_V_+1	96.1%	217	97.6%	63	0.5874	0.5802
gene216_I_1	84.9%	212	86.7%	64	0.6709	0.8958
gene216_G_-1	9.3%	210	9.5%	63	1.0000	0.9355
gene216_F_+1	65.2%	197	66.7%	57	0.8234	0.3665
gene216_F_1	96.8%	217	97.6%	62	0.7752	0.7715
gene216_D_1	0.0%	215	0.8%	64	0.2294	0.2294
gene216_D_-2	0.7%	214	0.8%	63	1.0000	1.0000

**TABLE 16**

BHR						
UK population						
GENE_EXON	Frequencies	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	65.8%	139	74.0%	48	0.1635	0.1885
gene216_T_3	8.3%	139	9.0%	50	0.8352	0.6515
gene216_T_4	97.1%	140	98.0%	49	1.0000	1.0000
gene216_T_5	75.4%	140	81.3%	48	0.2641	0.3646
gene216_T_6	78.5%	137	79.4%	46	1.0000	0.9547
gene216_T_7	97.5%	138	98.0%	50	1.0000	1.0000
gene216_T_8	97.8%	137	98.0%	49	1.0000	1.0000
gene216_T_+1	86.4%	140	94.0%	50	<b>0.0454</b>	0.1307
gene216_T_+2	37.9%	136	44.7%	47	0.2715	0.4549
gene216_T_+4	25.2%	139	26.0%	50	0.8938	0.1153
gene216_R_+2	87.5%	140	86.0%	50	0.7290	0.6834
gene216_R_+1	86.9%	122	92.6%	47	0.1838	0.3875
gene216_R_2	89.6%	134	94.8%	48	0.1494	0.4752
gene216_R_1	13.2%	140	7.0%	50	0.1041	0.3226
gene216_QR_+7	79.5%	139	85.0%	50	0.2983	0.3872
gene216_QR_+6	0.0%	139	0.0%	49	1.0000	1.0000
gene216_QR_+5	44.4%	133	49.0%	49	0.4771	0.5020
gene216_QR_+4	48.1%	128	57.3%	48	0.1508	0.2350
gene216_Q_+1	53.1%	129	48.9%	45	0.5407	0.6988
gene216_Q_2	72.9%	140	87.0%	50	<b>0.0038</b>	<b>0.0128</b>
gene216_Q_1	89.4%	132	95.8%	48	0.0613	0.1924
gene216_U_-1	86.1%	140	93.0%	50	0.0752	0.2087
gene216_L_+1	87.0%	138	94.0%	50	0.0638	0.2367
gene216_L_1	0.7%	140	1.0%	50	1.0000	1.0000
gene216_L_-1	87.2%	137	93.0%	50	0.1400	0.3796
gene216_L_-2	7.3%	137	9.0%	50	0.6623	0.5686
gene216_V_+2	71.6%	139	79.0%	50	0.1860	0.3615
gene216_V_+1	97.1%	140	98.0%	49	1.0000	1.0000
gene216_I_1	83.7%	138	91.0%	50	0.0952	0.2406
gene216_G_-1	9.9%	137	10.2%	49	1.0000	0.9269
gene216_F_+1	64.1%	128	73.3%	43	0.1466	0.2885
gene216_F_1	97.9%	140	97.9%	48	1.0000	1.0000
gene216_D_1	0.0%	139	1.0%	50	0.2646	0.2646
gene216_D_-2	0.7%	139	1.0%	49	1.0000	1.0000

**TABLE 17**

BHR						
US population						
GENE_EXON	Frequencies CNTL	N	CASE	N	ALLELE P-VALUE	GENOTYPE P-VALUE
gene216_T_2	67.8%	76	46.4%	14	0.0514	<b>0.0409</b>
gene216_T_3	9.5%	74	10.7%	14	0.7369	1.0000
gene216_T_4	94.7%	75	100.0%	13	0.6065	0.5986
gene216_T_5	79.2%	77	75.0%	14	0.6206	0.6767
gene216_T_6	76.6%	77	75.0%	14	0.8130	0.7738
gene216_T_7	94.2%	77	96.4%	14	1.0000	1.0000
gene216_T_8	93.9%	74	96.4%	14	1.0000	1.0000
gene216_T_+1	82.9%	76	78.6%	14	0.5937	0.6635
gene216_T_+2	36.3%	73	32.1%	14	0.8300	1.0000
gene216_T_+4	23.0%	76	28.6%	14	0.6296	0.7242
gene216_R_+2	89.6%	77	96.4%	14	0.4778	0.4545
gene216_R_+1	92.0%	69	76.9%	13	<b>0.0321</b>	<b>0.0452</b>
gene216_R_2	92.6%	74	78.6%	14	<b>0.0333</b>	<b>0.0469</b>
gene216_R_1	7.8%	77	28.6%	14	<b>0.0041</b>	<b>0.0072</b>
gene216_QR_+7	75.7%	76	71.4%	14	0.6391	0.2476
gene216_QR_+6	98.7%	77	100.0%	14	1.0000	1.0000
gene216_QR_+5	50.0%	77	39.3%	14	0.3130	0.4007
gene216_QR_+4	57.1%	77	64.3%	14	0.5371	0.8691
gene216_Q_+1	48.1%	77	62.5%	12	0.2724	0.4060
gene216_Q_2	75.3%	77	53.6%	14	<b>0.0233</b>	<b>0.0331</b>
gene216_Q_1	89.6%	77	87.5%	12	0.7250	0.5718
gene216_U_-1	83.1%	77	78.6%	14	0.5910	0.6593
gene216_L_+1	92.0%	75	75.0%	14	<b>0.0149</b>	<b>0.0227</b>
gene216_L_1	0.6%	77	0.0%	14	1.0000	1.0000
gene216_L_-1	92.0%	75	75.0%	14	<b>0.0149</b>	<b>0.0227</b>
gene216_L_-2	6.7%	75	7.1%	14	1.0000	1.0000
gene216_V_+2	70.8%	77	60.7%	14	0.3730	0.2711
gene216_V_+1	94.2%	77	96.4%	14	1.0000	1.0000
gene216_I_1	87.2%	74	71.4%	14	<b>0.0455</b>	<b>0.0463</b>
gene216_G_-1	8.2%	73	7.1%	14	1.0000	1.0000
gene216_F_+1	67.4%	69	46.4%	14	0.0510	0.0665
gene216_F_1	94.8%	77	96.4%	14	1.0000	1.0000
gene216_D_1	0.0%	76	0.0%	14	1.0000	1.0000
gene216_D_-2	0.7%	75	0.0%	14	1.0000	1.0000

**EXAMPLE 13: Haplotype analyses**

In addition to the analysis of individual SNPs, haplotype frequencies between the case and control groups were also compared. The haplotypes were constructed using a maximum likelihood approach. Since existing software for predicting haplotypes is unable to utilize individuals with missing data, a program was developed to make use of all individuals and, hence, provide more accurate haplotype frequency estimates. Haplotype analysis based on multiple SNPs in a gene is expected to provide increased evidence for an association between a given phenotype and that gene if all haplotyped SNPs are involved in the characterization of the phenotype. In other words, allelic variation involving those haplotyped SNPs are expected to be associated with different risks or susceptibilities toward the phenotype.

1. Asthma phenotype: The estimated frequency of each haplotype was compared between cases and controls by a permutation test. An overall comparison of the distribution of all haplotypes between the two groups was also performed. In Tables 18, 19 and 20 the haplotype analysis (2-at-a-time) for all SNPs in Gene 216 is presented for the combined, the UK and the US populations, respectively. The diagonal entries represent the single SNP p-values, while the other entries are the p-values for a test of association between the asthma phenotype and the haplotypes defined by the 2 SNPs listed on the horizontal and vertical axes. The frequency of the individual SNPs in the cases and controls are shown at the bottom of the tables. Colored cells indicate p-values that were statistically significant (light gray: 0.01 to 0.05, dark gray: 0.001 to 0.0099, black: < 0.001). As seen in Table 18, haplotypes defined by SNPs T5 & T8, SNPs T+2 & QR+4, T5 & T7 and SNPs T4 & T5, yielded highly significant p-values of 0.00039, 0.000042, 0.00056 and 0.00042 respectively, which were more significant than the analysis of these SNPs alone (T4 p = 0.16; T5 p = 0.04; T7 p = 0.16; T8 p = 0.25; T+2 p = 0.68; QR+4 p = 0.04). These associations were also more significant than the one observed for the single SNP T+1 reported above. In the UK population, the most significant association was found in Gene 216 (Table 19) with five

haplotypes significant at the 0.001 level (SNPs T+2 & QR+4,  $p = 0.000021$ ; QR+5 & QR+4,  $p = 0.00051$ ; QR+4 & Q+1  $p = 0.00066$ ; QR+6 & Q2,  $p = 0.00062$ ; and QR+4 & Q2,  $p = 0.00023$ ). Forty four haplotypes were significant at the 0.01 level in Gene 216 (Table 19) in the UK population. In the US  
5 population, numerous haplotypes were significant at the 0.01 level for Genes 216 (Table 20).

TABLE 18

	216_I_2	216_I_3	216_I_4	216_I_5	216_I_6	216_I_7	216_I_8	216_I_+1	216_I_+2	216_I_+4	216_I_+2	216_R_+2	216_R_+1	216_R_2	216_R_1	216_QR_+7	216_QR_+6	216_QR_+5	216_QR_+4
216_I_2	0.2029	0.1946	0.2058	0.156	0.4549	0.2035	0.9235	0.0244	0.374	0.3883	0.5729	0.543	0.543	0.1815	0.474	0.0826	0.4435	0.1096	0.0693
216_I_3		0.7841	0.1644	0.1063	0.7739	0.2105	0.2982	0.0203	0.9572	0.8651	0.839	0.9032	0.9032	0.5203	0.359	0.0254	0.6991	0.6339	0.0928
216_I_4			0.1576	0.00042	0.3636	0.3599	0.339	0.02	0.2702	0.2966	0.1468	0.2285	0.2285	0.2167	0.2531	0.0438	0.2735	0.4022	0.0632
216_I_5				0.042	0.0688	0.00056	0.00039	0.0011	0.1858	0.1501	0.0953	0.1449	0.1449	0.1286	0.1723	0.0529	0.2158	0.0562	0.0055
216_I_6					0.9235	0.2409	0.4041	0.0413	0.6676	0.7511	0.9381	0.8766	0.8766	0.8349	0.9771	0.0731	0.9419	0.2343	0.0468
216_I_7						0.1576	0.1239	0.0148	0.2381	0.2236	0.1581	0.2392	0.2392	0.2259	0.2448	0.0401	0.4254	0.4496	0.0853
216_I_8							0.2528	0.0156	0.4617	0.4724	0.3637	0.4603	0.4603	0.405	0.4688	0.0401	0.4254	0.4496	0.0853
216_I_+1								0.0055	0.6925	0.0181	0.0099	0.0158	0.0158	0.0216	0.0208	0.0195	0.0238	0.0193	0.00042
216_I_+2									0.6925	0.8136	0.8302	0.1038	0.1038	0.2308	0.2308	0.0232	0.7654	0.4372	0.00042
216_I_+4										0.5886	0.8604	0.8604	0.821	0.894	0.8173	0.0232	0.6812	0.5973	0.0053
216_R_+2											0.8076	0.8076	0.821	0.821	0.9409	0.0323	0.8444	0.8182	0.0666
216_R_+1														0.6372	0.6222	0.0473	0.9791	0.8605	0.0165
216_QR_+7														0.5928	0.8949	0.0598	0.7069	0.5374	0.0447
216_QR_+6															0.9025	0.0603	0.8982	0.2983	0.0449
216_QR_+5																0.016	0.0496	0.0102	0.0051
216_QR_+4																	0.6323	0.7003	0.1249
216_Q_1																		0.5794	0.0497
216_Q_2																			0.0387
216_Q_+1																			
216_Q_+2																			
216_Q_+4																			
216_Q_+5																			
216_Q_+6																			
216_Q_+7																			
216_U_+1																			
216_U_+2																			
216_U_+4																			
216_U_+5																			
216_U_+6																			
216_U_+7																			
216_L_+1																			
216_L_+2																			
216_L_+4																			
216_L_+5																			
216_L_+6																			
216_L_+7																			
216_V_+1																			
216_V_+2																			
216_V_+4																			
216_V_+5																			
216_V_+6																			
216_V_+7																			
216_G_+1																			
216_F_+1																			
216_F_+2																			
216_D_+1																			
216_D_+2																			
CNTL	66.50%	8.70%	96.30%	76.70%	77.80%	96.30%	96.50%	85.20%	37.30%	24.40%	88.30%	88.70%	88.80%	9.40%	11.30%	78.10%	0.50%	46.40%	51.50%
CASE	71.50%	9.50%	98.50%	83.30%	78.40%	98.50%	98.10%	92.40%	39.00%	26.30%	88.90%	88.80%	88.80%	10.80%	11.80%	85.70%	0.80%	48.80%	59.90%

TABLE 18 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_L_+1	216_L_-1	216_L_-2	216_V_+2	216_V_+1	216_L_1	216_G_-1	216_F_+1	216_F_1	216_D_1	216_D_-2
216_I_2	0.0775	0.2375	0.1151	0.0818	0.4173	0.6028	0.3661	0.2018	0.4952	0.4252	0.5976	0.4529	0.5379	0.1317	0.3817/216_I_2
216_I_3	0.0932	0.0511	0.0515	0.0574	0.2518	0.6065	0.5965	0.2331	0.3183	0.3637	0.8157	0.261	0.4089	0.4333	0.9258/216_I_3
216_I_4	0.307	0.092	0.0147	0.063	0.2503	0.2352	0.3023	0.1003	0.4414	0.3637	0.3654	0.2259	0.1664	0.064	0.2659/216_I_4
216_I_5	0.1526	0.0816	0.1052	0.0051	0.1666	0.1646	0.1146	0.0609	0.0073	0.1449	0.239	0.1489	0.0164	0.0252	0.1552/216_I_5
216_I_6	0.9167	0.1862	0.044	0.1048	0.9885	0.9852	0.9859	0.1586	0.7482	0.9858	0.9741	0.5124	0.6764	0.532	0.9411/216_I_6
216_I_7	0.3334	0.0671	0.0154	0.0423	0.2399	0.2367	0.2636	0.101	0.3118	0.3589	0.3635	0.1307	0.0162	0.0694	0.2707/216_I_7
216_I_8	0.5472	0.0961	0.0305	0.051	0.4685	0.4466	0.4668	0.2146	0.4267	0.6028	0.6774	0.2304	0.0388	0.1513	0.3915/216_I_8
216_I_+1	0.0236	0.0188	0.0168	0.0052	0.0177	0.0155	0.0161	0.063	0.0347	0.0366	0.0302	0.0254	0.0029	0.004	0.0228/216_I_+1
216_I_+2	0.9755	0.243	0.0533	0.0617	0.5266	0.5169	0.1993	0.0427	0.7458	0.9909	0.2258	0.3287	0.9157	0.4278	0.8918/216_I_+2
216_I_+4	0.8364	0.1748	0.0573	0.0924	0.8431	0.8582	0.3054	0.0849	0.8001	0.9064	0.421	0.2652	0.9665	0.3754	0.8453/216_I_+4
216_R_+2	0.8907	0.0886	0.0278	0.1086	0.9564	0.9531	0.8269	0.1872	0.705	0.9459	0.982	0.5521	0.9653	0.4631	0.8776/216_R_+2
216_R_+1	0.8471	0.1123	0.0448	0.0497	0.8427	0.8898	0.9227	0.3139	0.7802	0.8579	0.9953	0.426	0.9996	0.5725	0.9732/216_R_+1
216_R_2	0.8914	0.061	0.0589	0.0646	0.9528	0.8964	0.7497	0.8555	0.6835	0.461	0.9581	0.3036	0.8597	0.3897	0.831/216_R_2
216_R_1	0.8055	0.0543	0.0612	0.0399	0.8803	0.844	0.6875	0.2543	0.7844	0.9884	0.9417	0.372	0.9825	0.5493	0.9469/216_R_1
216_QR_+7	0.0096	0.046	0.0421	0.0208	0.0506	0.0531	0.7823	0.5044	0.6462	0.9622	0.8863	0.0792	0.023	0.0123	0.0599/216_QR_+7
216_QR_+6	0.8136	0.2301	0.0896	0.0915	0.9711	0.7055	0.8208	0.8586	0.2501	0.8781	0.205	0.4564	0.9706	0.5073	0.8973/216_QR_+6
216_QR_+5	0.418	0.0287	0.0255	0.0352	0.8441	0.6629	0.8208	0.076	0.2501	0.8781	0.8863	0.1319	0.1487	0.3787	0.8623/216_QR_+5
216_Q_+1	0.8075	0.0188	0.0494	0.0239	0.8011	0.8284	0.0418	0.0073	0.1122	0.189	0.205	0.0501	0.1038	0.0358	0.1518/216_Q_+1
216_Q_2	0.0432	0.0432	0.0517	0.052	0.0405	0.1089	0.9333	0.0127	0.8993	0.8755	0.4685	0.0448	0.7675	0.4463	0.9288/216_Q_2
216_Q_1	0.0213	0.0517	0.0517	0.052	0.0537	0.0553	0.0519	0.036	0.8993	0.8755	0.4685	0.0448	0.7675	0.4463	0.9288/216_Q_1
216_U_-1	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184	0.0184/216_U_-1
216_L_+1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_L_+1
216_L_1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_L_1
216_L_-1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_L_-1
216_L_-2	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_L_-2
216_V_+2	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_V_+2
216_V_+1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_V_+1
216_L_1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_L_1
216_G_-1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_G_-1
216_F_+1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_F_+1
216_F_1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_F_1
216_D_1	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_D_1
216_D_-2	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177	0.0177/216_D_-2
CNTL	51.20%	73.70%	89.50%	85.00%	88.70%	88.90%	92.90%	71.30%	96.10%	84.90%	90.70%	65.20%	96.80%	0.00%	0.70% CNTL
CASE	52.50%	80.50%	94.80%	91.20%	88.90%	89.20%	93.10%	77.10%	97.20%	85.30%	91.30%	70.40%	96.90%	0.40%	0.80% CASE

TABLE 19

	216 T 2	216 T 3	216 T 4	216 T 5	216 T 6	216 T 7	216 T 8	216 T +1	216 T +2	216 T +4	216 R +2	216 R +1	216 R 2	216 R 1	216 QR +7	216 QR +6	216 QR +5	216 QR +4
216 I 2	0.0566	0.0285	0.1466	0.0305	0.1853	0.1307	0.1754	0.0383	0.0323	0.0585	0.2643	0.1366	0.1597	0.1303	0.1229	0.0276	0.0383	0.0364
216 I 3		0.6308	0.5532	0.1098	0.851	0.4824	0.8448	0.0308	0.8738	0.7863	0.7026	0.4704	0.6808	0.2463	0.0778	0.214	0.3834	0.1567
216 I 4			0.3689	0.0182	0.5492	0.3627	0.3427	0.0275	0.6622	0.5907	0.4364	0.221	0.4296	0.1883	0.124	0.1116	0.3598	0.0605
216 I 5				0.0426	0.0384	0.0321	0.0224	0.035	0.0851	0.0784	0.0564	0.0256	0.0626	0.0246	0.0602	0.0173	0.0208	0.0024
216 I 6					0.7301	0.4807	0.7599	0.0624	0.7585	0.6073	0.5848	0.3842	0.8761	0.2345	0.207	0.2291	0.7201	0.0561
216 I 7						0.3129	0.2742	0.0314	0.489	0.4026	0.241	0.1683	0.3089	0.1454	0.169	0.0765	0.2753	0.0444
216 I 8							0.7388	0.0255	0.814	0.9112	0.5594	0.3237	0.597	0.2337	0.1538	0.2448	0.5088	0.073
216 I +1								0.0105	0.0326	0.0257	0.0387	0.0064	0.02	0.0032	0.0292	0.0031	0.0145	0.000321
216 I +2									0.5682	0.9467	0.8842	0.104	0.3214	0.2057	0.0082	0.1953	0.1004	0.000321
216 I +4										0.9163	0.987	0.4987	0.7548	0.3174	0.0654	0.26	0.2475	0.0102
216 R +2												0.2758	0.7381	0.2765	0.1148	0.2761	0.441	0.0344
216 R +4												0.2211	0.3444	0.2566	0.0357	0.075	0.1305	0.0032
216 R +5													0.4279	0.2905	0.0876	0.1875	0.2629	0.0237
216 Q +1														0.1473	0.0287	0.059	0.1383	0.0037
216 Q 2															0.0535	0.026	0.0095	0.007
216 Q +1																0.1806	0.1177	0.0124
216 U +1																	0.2273	0.00351
216 L +1																		0.0229
216 L -1																		
216 L -2																		
216 V +2																		
216 V +1																		
216 Q -1																		
216 F +1																		
216 F -1																		
216 D -1																		
216 D -2	65.80%	8.30%	97.10%	75.40%	78.50%	97.50%	97.80%	86.40%	37.90%	25.20%	87.50%	86.30%	10.50%	13.20%	79.50%	0.00%	44.40%	48.10%
CNTL	74.30%	9.60%	96.50%	63.30%	80.10%	99.00%	98.50%	93.30%	40.50%	25.00%	87.50%	91.10%	8.20%	8.70%	86.40%	1.00%	50.00%	59.10%
CASE																		



TABLE 19 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_L_+1	216_L_1	216_L_-1	216_L_-2	216_V_-2	216_V_+1	216_L_1	216_G_-1	216_F_+1	216_F_1	216_D_1	216_D_-2	
216_T_2	0.0594	0.0177	0.0595	0.0766	0.13	0.1845	0.2392	0.069	0.3132	0.2304	0.2111	0.0772	0.182	0.0447	0.1959	216_T_2	
216_T_3	0.7224	0.003	0.0764	0.0983	0.1526	0.6246	0.3066	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_3	
216_T_4	0.5984	0.0063	0.0595	0.0913	0.14	0.5646	0.1154	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_4	
216_T_5	0.1084	0.0098	0.0825	0.0661	0.0192	0.173	0.186	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_5	
216_T_6	0.8693	0.0149	0.0602	0.1557	0.2168	0.5078	0.196	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_6	
216_T_7	0.5518	0.0095	0.0537	0.082	0.1031	0.4799	0.0878	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_7	
216_T_8	0.5879	0.0097	0.0927	0.0985	0.1965	0.8351	0.1832	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_8	
216_T_+1	0.0185	0.0053	0.0305	0.0152	0.0044	0.0305	0.0032	0.0381	0.0563	0.0285	0.0174	0.0428	0.0141	0.0235	0.0339	216_T_+1	
216_T_+2	0.3888	0.0173	0.0816	0.0518	0.2532	0.9068	0.2357	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_+2	
216_T_+4	0.8036	0.0116	0.0791	0.1923	0.2638	0.5942	0.2499	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_T_+4	
216_R_+2	0.8431	0.0039	0.0365	0.0053	0.2322	0.999	0.2051	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_R_+2	
216_R_+1	0.3291	0.0178	0.0193	0.02	0.3412	0.4162	0.142	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_R_+1	
216_R_2	0.6168	0.0149	0.0553	0.0626	0.2429	0.8853	0.3641	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_R_2	
216_R_1	0.3033	0.0168	0.0183	0.0181	0.3907	0.8686	0.3554	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_R_1	
216_QR_+7	0.014	0.0176	0.0816	0.039	0.0233	0.1904	0.0154	0.0213	0.1272	0.1184	0.0743	0.1161	0.0394	0.1353	0.0429	216_QR_+7	
216_QR_+6	0.244	0.0062	0.0092	0.012	0.0464	0.2566	0.0462	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_QR_+6	
216_QR_+5	0.2521	0.0072	0.0162	0.0239	0.1056	0.4597	0.0951	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_QR_+5	
216_QR_+4	0.00036	0.00023	0.0045	0.004	0.0073	0.0551	0.0063	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_QR_+4	
216_Q_+1	0.6345	0.0075	0.0393	0.0175	0.2369	0.8197	0.1096	0.069	0.1845	0.2392	0.2111	0.0772	0.182	0.0447	0.1959	216_Q_+1	
216_Q_2	0.002	0.0049	0.0075	0.0075	0.0075	0.0042	0.003	0.0071	0.0053	0.0072	0.0107	0.0181	0.0043	0.0077	0.0013	216_Q_2	
216_Q_1	0.0274	0.00629	0.0182	0.0182	0.0182	0.1052	0.0127	0.0826	0.0403	0.0896	0.0126	0.0922	0.0242	0.1105	0.0232	216_Q_1	
216_U_-1	0.0419	0.0137	0.0137	0.0137	0.0137	0.1417	0.0126	0.1019	0.0759	0.084	0.0164	0.2182	0.0304	0.0805	0.0268	216_U_-1	
216_L_+1	0.1059	0.1059	0.1059	0.1059	0.1059	0.3583	0.134	0.2587	0.0166	0.209	0.3181	0.1408	0.1	0.1895	0.0907	216_L_+1	
216_L_1	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	0.334	216_L_1	
216_L_-1	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	0.0892	216_L_-1	
216_L_-2	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	216_L_-2	
216_V_+2	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_V_+2	
216_V_+1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_V_+1	
216_L_1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_L_1	
216_G_1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_G_1	
216_F_+1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_F_+1	
216_F_1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_F_1	
216_D_1	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_D_1	
216_D_-2	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717	216_D_-2	
CNTL	53.10%	72.90%	59.40%	86.10%	87.00%	99.30%	87.20%	92.70%	71.60%	97.10%	88.70%	90.20%	64.10%	97.90%	0.00%	7.00% CNTL	
CASE	50.50%	84.80%	95.10%	92.30%	91.80%	99.50%	92.20%	92.30%	79.10%	98.00%	89.20%	90.10%	74.20%	99.00%	1.00%	1.00% CASE	

**TABLE 20**

[illegible]

TABLE 20 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_L_+1	216_L_1	216_L_-1	216_L_2	216_V_-2	216_V_+1	216_L_1	216_G_-1	216_F_+1	216_F_-1	216_D_-1	216_D_-2
216_I_2	0.3352	0.3378	0.59	0.5138	0.0463	0.5287	0.0468	0.3069	0.7057	0.8171	0.0211	0.3718	0.3846	0.1695	0.4097	0.5097 216_I_2
216_I_3	0.3116	0.1334	0.4369	0.8853	0.0278	0.9279	0.0234	0.3854	0.9666	0.7054	0.0674	0.408	0.3626	0.4632	0.9982	0.9788 216_I_3
216_I_4	0.3128	0.2639	0.4547	0.799	0.0246	0.5377	0.0227	0.3488	0.7228	0.7844	0.0086	0.2655	0.373	0.4659	0.375	0.5428 216_I_4
216_I_5	0.3843	0.1539	0.8504	0.0324	0.0435	0.5676	0.041	0.0331	0.514	0.1163	0.0093	0.0509	0.0382	0.2078	0.4926	0.7133 216_I_5
216_I_6	0.4714	0.1906	0.6257	0.7187	0.0187	0.8442	0.0161	0.6166	0.834	0.8164	0.0025	0.4371	0.2211	0.7823	0.5572	0.6864 216_I_6
216_I_7	0.4482	0.1868	0.6435	0.8069	0.0433	0.8393	0.0428	0.5364	0.9562	0.8494	0.0064	0.4369	0.2891	0.1189	0.5898	0.7552 216_I_7
216_I_8	0.4398	0.1691	0.8502	0.7154	0.048	0.8725	0.0378	0.5089	0.9457	0.845	0.0087	0.4309	0.2412	0.1106	0.5399	0.8863 216_I_8
216_I_+1	0.8222	0.0954	0.6821	0.5049	0.0425	0.5648	0.0441	0.3063	0.8235	0.923	0.0509	0.4246	0.2456	0.1913	0.4553	0.625 216_I_+1
216_I_+2	0.1783	0.4184	0.8575	0.7884	0.0414	0.7988	0.0437	0.6796	0.8393	0.9393	0.0249	0.4126	0.5057	0.8372	0.7438	0.8385 216_I_+2
216_I_+4	0.3097	0.1352	0.6432	0.6772	0.0153	0.6708	0.0166	0.5517	0.8063	0.9142	0.0074	0.3463	0.1986	0.7054	0.5583	0.5808 216_I_+4
216_R_+2	0.0633	0.2517	0.3815	0.3853	0.0816	0.4186	0.0284	0.2798	0.6015	0.5281	0.0181	0.1911	0.3275	0.4783	0.2657	0.4241 216_R_+2
216_R_+1	0.0859	0.1495	0.095	0.0889	0.0385	0.0492	0.0842	0.0908	0.6032	0.0816	0.0484	0.0573	0.2097	0.0621	0.028	0.0416 216_R_+1
216_R_2	0.071	0.0779	0.0636	0.0636	0.0155	0.0383	0.0137	0.0955	0.6354	0.0637	0.0386	0.0503	0.0588	0.0501	0.0189	0.0363 216_R_2
216_R_1	0.0233	0.0188	0.028	0.0225	0.0052	0.009	0.0057	0.0178	0.0036	0.0176	0.0269	0.0112	0.0368	0.0036	0.009	0.0095 216_R_1
216_QR_+7	0.4567	0.0505	0.5493	0.477	0.0291	0.4631	0.0334	0.5076	0.7221	0.8151	0.0351	0.4658	0.0597	0.3088	0.2805	0.5 216_QR_+7
216_QR_+8	0.1885	0.265	0.7142	0.824	0.0213	0.8063	0.0208	0.7791	0.9583	0.7173	0.0128	0.6403	0.2129	0.5454	0.8023	0.6403 216_QR_+8
216_QR_+5	0.7691	0.0758	0.673	0.8434	0.0448	0.5902	0.0455	0.4553	0.6143	0.0827	0.0285	0.5385	0.4736	0.0228	0.441	0.6098 216_QR_+5
216_QR_+4	0.0526	0.4984	0.4435	0.7892	0.0355	0.6247	0.046	0.6027	0.2742	0.7616	0.0035	0.5038	0.451	0.5462	0.4629	0.6479 216_QR_+4
216_Q_+1	0.1345	0.0246	0.3075	0.6184	0.0426	0.2997	0.04	0.1554	0.3921	0.4884	0.0538	0.2419	0.1754	0.3913	0.1119	0.2253 216_Q_+1
216_Q_2	0.1571	0.1571	0.1131	0.0323	0.0518	0.3763	0.0483	0.3266	0.205	0.3191	0.0184	0.2751	0.2713	0.0371	0.1586	0.3726 216_Q_2
216_Q_1	0.5726	0.1571	0.1131	0.6694	0.0482	0.5976	0.0487	0.1405	0.579	0.6386	0.0309	0.1549	0.2302	0.5842	0.4461	0.5888 216_Q_1
216_U_-1	0.6654	0.6654	0.6632	0.6632	0.0475	0.6632	0.0416	0.3148	0.8367	0.9328	0.0434	0.2469	0.2752	0.1935	0.4873	0.6241 216_U_-1
216_L_+1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_L_+1
216_L_1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_L_1
216_L_-1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_L_-1
216_L_2	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_L_2
216_V_+2	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_V_+2
216_V_+1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_V_+1
216_G_-1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_G_-1
216_F_+1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_F_+1
216_F_1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_F_1
216_D_1	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_D_1
216_D_-2	0.0116	0.0116	0.0222	0.0116	0.0116	0.0222	0.0116	0.0444	0.02	0.0404	0.0287	0.0321	0.0504	0.0323	0.0126	0.0205 216_D_-2
CNTL	48.10%	75.30%	89.80%	83.10%	92.00%	99.40%	92.00%	93.30%	70.80%	94.20%	87.20%	91.80%	67.40%	94.80%	0.00%	0.70% CNTL
CASE	60.90%	64.80%	93.80%	87.00%	77.90%	100.00%	77.80%	96.30%	69.20%	94.20%	70.40%	86.20%	57.40%	92.60%	0.00%	0.00% CASE

2. Bronchial Hyper-responsiveness: A similar test for association of 2-SNP-a-time haplotypes with BHR ( $PC_{20} \leq 16$  mg/ml) was performed. In Tables 21, 22 and 23, the haplotype analysis (2-at-a-time) for all SNPs in Gene 216 is presented for the combined, the UK and the US populations, respectively. One haplotype in Gene 216 (Table 21: SNPs T+2 & QR+4,  $p=0.0041$ ) was significant at the 0.01 level in the combined sample. In contrast, in the UK population, seventeen haplotypes were significant at the 0.01 level in Gene 216 (Table 22). In the US population, nine haplotypes were significant at the 0.01 level in Gene 216 (Table 23). Tables 18, 19, and 20 and Tables 21, 22 and 23 showed similar patterns of significance with lower level achieved in the BHR analysis due to the reduced sample size in the ( $PC_{20} \leq 16$  mg/ml) subgroup.

In summary, haplotype analysis of SNPs significantly strengthened the evidence in support of Gene 216 as an asthma susceptibility gene. In some SNP combinations, the association was increased by an order of magnitude. The most striking association again appeared in the 3' region of the gene, in agreement with the single SNP analysis.

TABLE 21

	216_T_2	216_T_3	216_T_4	216_T_5	216_T_6	216_T_7	216_T_8	216_T_+1	216_T_+2	216_T_+4	216_R_+2	216_R_+1	216_R_2	216_R_1	216_QR_+7	216_QR_+6	216_QR_+5	216_QR_+4
216_T_2	0.8294	0.8699	0.8528	0.9001	0.6497	0.8157	0.8569	0.3137	0.8102	0.8854	0.9227	0.8548	0.9513	0.9968	0.6127	0.809	0.4618	0.3714
216_T_3		0.8592	0.4838	0.7833	0.8337	0.9421	0.8733	0.2704	0.4217	0.5951	0.9397	0.71	0.8294	0.9975	0.6953	0.8968	0.9682	0.5292
216_T_4			0.3878	0.0801	0.6596	0.9816	0.6683	0.4037	0.426	0.474	0.2036	0.4754	0.4846	0.4814	0.6119	0.2507	0.6758	0.372
216_T_5				0.5428	0.5556	0.0375	0.0409	0.0359	0.6617	0.7754	0.7408	0.8658	0.7605	0.8568	0.7124	0.6804	0.6199	0.2883
216_T_6					1	0.7439	0.7538	0.3946	0.7367	0.6602	0.866	0.8322	0.9808	0.9949	0.5552	0.8049	0.937	0.2731
216_T_7						0.5856	0.3561	0.264	0.5599	0.7258	0.4951	0.7442	0.746	0.7652	0.5427	0.4354	0.8044	0.33
216_T_8							0.7756	0.2735	0.614	0.7615	0.5474	0.7837	0.7875	0.795	0.5818	0.4673	0.8358	0.3389
216_T_+1								0.1413	0.4229	0.2936	0.1518	0.1374	0.2295	0.1848	0.2476	0.2972	0.2684	0.1303
216_T_+2									0.3978	0.8361	0.6592	0.1325	0.3537	0.3367	0.1168	0.4177	0.8063	0.0324
216_T_+4										0.8421	0.9199	0.6872	0.8959	0.8584	0.4637	0.6411	0.827	0.0709
216_R_+2											1	0.8319	0.9837	0.9783	0.4454	0.8445	0.9937	0.3199
216_R_+1												1	0.8591	0.5707	0.6045	0.7804	0.9553	0.0906
216_QR_+2													1	0.338	0.6225	0.7814	0.977	0.208
216_QR_+1														0.875	0.6402	0.8153	0.3556	0.197
216_Q_2															0.3876	0.3756	0.0564	0.1251
216_Q_1																1	0.8309	0.21
216_U_+1																	1	0.0265
216_L_+1																		0.1521
216_L_1																		
216_L_+2																		
216_V_+1																		
216_L_1																		
216_Q_1																		
216_F_+1																		
216_F_1																		
216_D_1																		
216_D_2																		
CNTL	66.50%	8.70%	96.30%	76.70%	77.80%	96.30%	96.50%	85.20%	37.30%	24.40%	88.30%	88.70%	90.80%	11.30%	76.10%	99.50%	48.40%	51.50%
CASE	67.70%	9.40%	98.40%	79.80%	78.30%	97.70%	97.60%	90.60%	41.80%	26.60%	88.30%	89.20%	91.10%	11.70%	82.00%	100.00%	46.80%	58.90%

TABLE 21 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_1	216_L_+1	216_L_1	216_L_-1	216_L_-2	216_V_+2	216_V_+1	216_L_1	216_G_-1	216_F_+1	216_F_1	216_D_1	216_D_-2	
216_I_2	0.2027	0.4995	0.5707	0.6007	0.9369	0.9292	0.9947	0.7441	0.8759	0.7461	0.9473	0.9753	0.9665	0.9031	0.2822	0.9148 216_I_2	
216_I_3	0.9662	0.6224	0.1274	0.3869	0.9807	0.9266	0.9835	0.6939	0.6655	0.5785	0.8079	0.8659	0.894	0.732	0.2896	0.9415 216_I_3	
216_I_4	0.5948	0.4522	0.1688	0.4878	0.4333	0.6387	0.4707	0.5001	0.3621	0.6522	0.6115	0.6035	0.6912	0.5512	0.1272	0.6373 216_I_4	
216_I_5	0.7883	0.5364	0.4421	0.0178	0.8419	0.8272	0.8799	0.7291	0.5465	0.0255	0.3726	0.9338	0.8212	0.0255	0.1684	0.7477 216_I_5	
216_I_6	0.9839	0.431	0.2757	0.5172	0.9378	0.3689	0.9918	0.6763	0.6743	0.7197	0.4388	0.9627	0.6782	0.9025	0.3121	0.9483 216_I_6	
216_I_7	0.9655	0.374	0.2662	0.3785	0.6817	0.774	0.7393	0.7205	0.6851	0.7075	0.8106	0.8494	0.749	0.4337	0.1823	0.7718 216_I_7	
216_I_8	0.901	0.3728	0.3041	0.3782	0.7262	0.7847	0.7922	0.7622	0.7673	0.6937	0.6163	0.9027	0.7612	0.539	0.1647	0.7718 216_I_8	
216_I_+1	0.349	0.2692	0.2569	0.1884	0.1217	0.2688	0.1549	0.2174	0.3823	0.5451	0.6303	0.4265	0.4195	0.2726	0.0542	0.7493 216_I_+1	
216_I_+2	0.7037	0.5357	0.3112	0.5279	0.8787	0.477	0.8913	0.3029	0.4624	0.5545	0.7345	0.2619	0.4868	0.6347	0.3025	0.8406 216_I_+2	
216_I_+4	0.9869	0.3483	0.1773	0.0904	0.9371	0.9874	0.9912	0.9466	0.694	0.5058	0.8759	0.999	0.9922	0.9141	0.3019	0.9733 216_I_+4	
216_R_+1	0.8743	0.5514	0.1027	0.2215	0.8563	0.9592	0.5874	0.8544	0.6935	0.7064	0.6623	0.8918	0.7704	0.9203	0.3276	0.8734 216_R_+1	
216_R_2	0.9485	0.4841	0.1249	0.3228	0.9387	0.9506	0.2671	0.8658	0.7229	0.7163	0.7775	0.9891	0.9096	0.9257	0.2969	0.94 216_R_2	
216_R_1	0.3474	0.2054	0.1515	0.2689	0.5129	0.724	0.6153	0.7354	0.6436	0.5662	0.5829	0.5411	0.7774	0.6443	0.1341	0.7947 216_R_1	
216_QR_+7	0.0222	0.497	0.3994	0.3418	0.5129	0.9251	0.9328	0.8535	0.0245	0.3431	0.3571	0.9664	0.7774	0.6443	0.1341	0.81 216_QR_+7	
216_QR_+6	0.8314	0.3809	0.3646	0.5159	0.9356	0.977	0.2002	0.2075	0.0146	0.8169	0.9418	0.9434	0.5968	0.355	0.087	0.81 216_QR_+6	
216_QR_+5	0.9882	0.3183	0.2453	0.185	0.9356	0.977	0.3135	0.8638	0.3227	0.4849	0.4143	0.5794	0.4052	0.9075	0.2945	0.9019 216_QR_+5	
216_QR_+4	0.0611	0.1626	0.1368	0.1257	0.9356	0.977	0.1721	0.3748	0.3227	0.4849	0.4143	0.5794	0.4052	0.9075	0.2945	0.4236 216_QR_+4	
216_Q_+1	1	0.2097	0.2725	0.3284	0.3249	0.5827	0.2002	0.2075	0.3227	0.4849	0.4143	0.5794	0.4052	0.9075	0.2945	0.8972 216_Q_+1	
216_Q_2	0.2009	0.1565	0.3331	0.3747	0.1173	0.3107	0.1415	0.1125	0.268	0.1866	0.135	0.3023	0.3737	0.2772	0.0535	0.6668 216_Q_2	
216_Q_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.311 216_Q_1	
216_U_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5417 216_U_1	
216_U_+1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8878 216_U_+1	
216_U_+2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8815 216_U_+2	
216_U_+3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.6135 216_U_+3	
216_U_+4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8115 216_U_+4	
216_V_+1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.7397 216_V_+1	
216_V_+2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.7527 216_V_+2	
216_V_+3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8588 216_V_+3	
216_V_+4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9888 216_V_+4	
216_G_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9023 216_G_1	
216_F_+1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9023 216_F_+1	
216_F_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9023 216_F_1	
216_D_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.2138 216_D_1	
216_D_2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.70% CNTL	
CASE	51.80%	73.70%	89.50%	85.00%	88.70%	70%	89.10%	7.10%	71.30%	96.10%	84.90%	9.30%	65.20%	96.80%	0.00%	0.70% CASE	
	51.80%	79.70%	94.20%	89.80%	88.80%	0.80%	89.10%	8.60%	75.00%	97.60%	86.70%	9.50%	66.70%	97.60%	0.80%		

**TABLE 22**

	216_I_2	216_I_3	216_I_4	216_I_5	216_I_6	216_I_7	216_I_8	216_I_+1	216_I_+2	216_I_+4	216_I_+2	216_I_+1	216_I_2	216_I_2	216_R_1	216_QR_+7	216_QR_+6	216_QR_+5	216_QR_+4
216_I_2	0.1835	0.1946	0.4642	0.2542	0.2634	0.4757	0.4874	0.1699	0.0577	0.2963	0.3235	0.2664	0.1937	0.2668	0.41	0.151	0.0598	0.0412	
216_I_3		0.8352	0.9346	0.5085	0.9737	0.9445	0.9524	0.1119	0.2317	0.5899	0.8441	0.4145	0.2901	0.4045	0.0968	0.8146	0.7043	0.1885	
216_I_4			1	0.1832	0.891	0.7313	0.4727	0.1102	0.6516	0.8303	0.5016	0.279	0.2638	0.2006	0.4975	0.7313	0.7734	0.3353	
216_I_5				0.2641	0.2802	0.221	0.2228	0.0834	0.3727	0.415	0.2666	0.1992	0.132	0.15	0.3394	0.2043	0.1396	0.0658	
216_I_6					1	0.5555	0.3371	0.2145	0.6993	0.4986	0.6354	0.3337	0.3054	0.2174	0.8461	0.7353	0.2584		
216_I_7						1	0.7135	0.105	0.6713	0.7568	0.5668	0.277	0.2847	0.235	0.4779	0.8161	0.7714	0.3188	
216_I_8							1	0.0869	0.6249	0.8535	0.5841	0.3144	0.3013	0.2432	0.4423	0.8072	0.8158	0.3265	
216_I_+1								0.0454	0.0696	0.6143	0.0171	0.0159	0.0181	0.0177	0.0593	0.0403	0.0199	0.0315	
216_I_+2									0.2715	0.8938	0.6109	0.1093	0.1696	0.1672	0.0291	0.2774	0.0937	0.0027	
216_I_+4										0.8941	0.8941	0.4271	0.3497	0.2603	0.3809	0.8728	0.5997	0.1248	
216_R_+1											0.729	0.2922	0.3083	0.2672	0.3422	0.6525	0.5864	0.1828	
216_R_2												0.1838	0.3156	0.2629	0.1076	0.1376	0.2608	0.0485	
216_R_1													0.1494	0.1886	0.1264	0.1401	0.2309	0.0603	
216_QR_+7														0.1041	0.0702	0.1153	0.2645	0.0454	
216_QR_+6															0.2383	0.2237	0.0061	0.0362	
216_QR_+5																1	0.4576	0.0023	
216_QR_+4																	0.4771	0.1508	
216_Q_+1																			
216_Q_2																			
216_Q_1																			
216_U_+1																			
216_L_+1																			
216_L_1																			
216_L_+1																			
216_L_2																			
216_V_+2																			
216_V_+1																			
216_L_1																			
216_G_+1																			
216_F_+1																			
216_F_1																			
216_D_1																			
216_D_2																			
CNTL	65.80%	8.30%	97.10%	75.40%	78.50%	97.50%	97.80%	86.40%	37.80%	25.20%	87.50%	86.90%	89.60%	13.20%	79.50%	0.00%	44.40%	48.10%	
CASE	74.00%	9.00%	98.00%	81.30%	79.40%	98.00%	98.00%	94.00%	44.70%	26.00%	85.00%	92.60%	94.80%	7.00%	85.00%	0.00%	49.00%	57.30%	

TABLE 22 (CONT')

	216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_L_+1	216_L_1	216_L_-1	216_L_-2	216_V_+2	216_V_+1	216_L_1	216_G_-1	216_F_+1	216_F_1	216_D_1	216_D_-2	
216_T_2	0.0993	0.0239	0.1363	0.2354	0.2108	0.4205	0.2791	0.1304	0.2571	0.4658	0.3122	0.4106	0.2736	0.4814	0.0819	0.3287 216_T_2	
216_T_3	0.6845	0.0298	0.1044	0.1847	0.2611	0.9291	0.4768	0.4817	0.202	0.9277	0.3057	0.6906	0.1932	0.9356	0.3363	0.8504 216_T_3	
216_T_4	0.8254	0.0112	0.1712	0.1539	0.1427	0.8566	0.2395	0.7921	0.3405	0.7155	0.1905	0.6917	0.274	0.738	0.2708	0.8707 216_T_4	
216_T_5	0.3609	0.0391	0.1894	0.0872	0.097	0.7714	0.1855	0.507	0.3068	0.1748	0.1044	0.6102	0.3101	0.1393	0.1039	0.6496 216_T_5	
216_T_6	0.8493	0.0175	0.1382	0.3072	0.1424	0.4032	0.2802	0.9045	0.3762	0.9218	0.2811	0.9952	0.2941	0.9375	0.3535	0.9013 216_T_6	
216_T_7	0.8564	0.0067	0.2165	0.1593	0.1485	0.9358	0.2551	0.8976	0.3598	0.7347	0.1891	0.9602	0.2734	0.6747	0.3149	0.8617 216_T_7	
216_T_8	0.8823	0.0082	0.2404	0.1703	0.1848	0.9092	0.2773	0.9105	0.4828	0.476	0.1876	0.9762	0.2684	0.8999	0.3484	0.9021 216_T_8	
216_T_+1	0.0873	0.014	0.1085	0.1081	0.0033	0.0936	0.0178	0.1105	0.108	0.1116	0.0144	0.1974	0.0741	0.0772	0.0219	0.0995 216_T_+1	
216_T_+2	0.0472	0.0267	0.1205	0.1096	0.1576	0.0858	0.2184	0.4892	0.0177	0.6495	0.3096	0.3468	0.1153	0.6344	0.1181	0.6945 216_T_+2	
216_T_+4	0.4254	0.0175	0.1491	0.2888	0.1708	0.8009	0.3196	0.8842	0.1628	0.5625	0.2711	0.6529	0.2089	0.8494	0.3638	0.9405 216_T_+4	
216_R_+2	0.6558	0.0049	0.0434	0.0064	0.1525	0.9209	0.2969	0.7197	0.3294	0.5441	0.1899	0.9553	0.519	0.6061	0.267	0.8387 216_R_+2	
216_R_+1	0.2608	0.0322	0.0119	0.0317	0.2572	0.362	0.6439	0.4755	0.1176	0.2779	0.3641	0.5165	0.2261	0.3177	0.061	0.4847 216_R_+1	
216_R_2	0.2443	0.0239	0.0173	0.0817	0.2558	0.3122	0.2887	0.3547	0.0767	0.2614	0.1962	0.4256	0.2238	0.2998	0.0745	0.4318 216_R_2	
216_R_1	0.2559	0.0333	0.0103	0.0243	0.1921	0.2283	0.1027	0.2925	0.0432	0.2021	0.2913	0.3177	0.2156	0.2523	0.0423	0.2834 216_R_1	
216_OR_+7	0.0061	0.0423	0.1453	0.18	0.0323	0.6664	0.0796	0.125	0.3654	0.5079	0.1117	0.3534	0.2261	0.4609	0.1223	0.6626 216_OR_+7	
216_OR_+6	0.55	0.002	0.0455	0.0563	0.0568	0.8893	0.1137	0.65	0.1144	0.7358	0.0542	0.9403	0.1299	0.9983	0.2238	0.826 216_OR_+6	
216_OR_+5	0.8707	0.0113	0.0561	0.0036	0.1599	0.7718	0.0423	0.6116	0.0079	0.7782	0.2914	0.7551	0.1616	0.815	0.2255	0.8636 216_OR_+5	
216_OR_+4	0.0137	0.0029	0.0531	0.0259	0.0297	0.2615	0.0558	0.1816	0.0023	0.3334	0.1253	0.2598	0.0624	0.3336	0.0825	0.2851 216_OR_+4	
216_Q_+1	0.5407	0.0211	0.0536	0.062	0.1512	0.7981	0.0497	0.5841	0.0023	0.8282	0.3082	0.4634	0.1278	0.8845	0.2558	0.8262 216_Q_+1	
216_Q_2		0.0033	0.0125	0.0142	0.0101	0.0122	0.0106	0.0118	0.0098	0.0106	0.0101	0.037	0.021	0.0077	0.0017	0.0152 216_Q_2	
216_U_-1			0.0613	0.1043	0.0077	0.1176	0.0182	0.7099	0.094	0.1753	0.0129	0.1534	0.0824	0.2463	0.0268	0.1185 216_U_-1	
216_L_+1				0.0752	0.0112	0.1367	0.0283	0.1799	0.0883	0.1592	0.0097	0.3854	0.1051	0.1615	0.0283	0.1401 216_L_+1	
216_L_1					0.0638	0.1326	0.0572	0.1554	0.0261	0.1422	0.2061	0.1806	0.1294	0.1884	0.0302	0.1897 216_L_1	
216_L_-1							0.2934	0.8458	0.1022	0.8556	0.1542	0.9613	0.325	0.9999	0.3917	0.8772 216_L_-1	
216_L_-2					0.14		0.14	0.3116	0.0638	0.2334	0.2875	0.3511	0.2591	0.2871	0.0528	0.1734 216_L_-2	
216_V_+2								0.6823	0.2018	0.7842	0.2354	0.8657	0.1417	0.9987	0.2811	0.8408 216_V_+2	
216_V_+1									0.186	0.3583	0.0558	0.3687	0.2253	0.4145	0.077	0.4394 216_V_+1	
216_L_1										1	0.1857	0.8927	0.2713	0.7385	0.2762	0.867 216_L_1	
216_G_-1											0.0952	0.3777	0.1788	0.1961	0.0832	0.2355 216_G_-1	
216_F_+1												1	0.2373	0.9526	0.9634	0.982 216_F_+1	
216_F_1													0.1466	0.2595	0.0691	0.2676 216_F_1	
216_D_1														1	0.383	0.9241 216_D_1	
216_D_-2															0.2646	0.2646 216_D_-2	
CNTL	53.10%	72.90%	89.40%	86.10%	87.00%	0.70%	87.20%	7.30%	71.60%	97.10%	83.70%	9.90%	64.10%	97.90%	0.00%	0.70% CNTL	
CASE	46.90%	87.00%	95.80%	93.00%	94.00%	1.00%	93.00%	9.00%	79.00%	98.00%	91.00%	10.20%	73.30%	97.90%	1.00%	1.00% CASE	



TABLE 23

	216_T_2	216_T_3	216_T_4	216_T_5	216_T_6	216_T_7	216_T_8	216_T_+1	216_T_+2	216_T_+4	216_R_+2	216_R_+1	216_R_2	216_R_1	216_QR_+7	216_QR_+6	216_QR_+5	216_QR_+4
216_T_2	0.0514	0.1243	0.0357	0.0497	0.0787	0.0595	0.082	0.1864	0.2019	0.0488	0.1508	0.1085	0.0837	0.0083	0.0914	0.0506	0.1874	0.1924
216_T_3		0.7369	0.4787	0.4804	0.6341	0.8292	0.8018	0.9402	0.9479	0.9244	0.1641	0.165	0.1963	0.0202	0.2541	0.9542	0.5592	0.3088
216_T_4			0.6065	0.3323	0.5394	0.9824	0.9942	0.5099	0.3588	0.6893	0.1293	0.0426	0.0404	0.0087	0.509	0.2862	0.2802	0.5248
216_T_5				0.6206	0.6544	0.168	0.2128	0.1058	0.5476	0.9089	0.5923	0.1377	0.1245	0.0105	0.0193	0.6959	0.5418	0.8568
216_T_6					0.813	0.812	0.74	0.8376	0.7371	0.7783	0.3655	0.1368	0.1121	0.0189	0.576	0.7745	0.6551	0.5037
216_T_7						1	0.5041	0.6321	0.7986	0.7716	0.3798	0.1355	0.1386	0.0189	0.5855	0.6202	0.3781	0.7488
216_T_8							1	0.488	0.8346	0.6672	0.4505	0.1151	0.1286	0.0196	0.9187	0.6807	0.6064	0.9072
216_T_+1								0.5837	0.83	0.2834	0.4835	0.1588	0.1904	0.0325	0.9216	0.7751	0.4055	0.8702
216_T_+2										0.6236	0.4691	0.0794	0.1096	0.0098	0.8987	0.6591	0.6792	0.2344
216_T_+4											0.4778	0.1289	0.14	0.0136	0.2244	0.3101	0.4372	0.3444
216_T_+6												0.0921	0.1017	0.011	0.0845	0.0921	0.1826	0.1505
216_T_+8												0.0321	0.0333	0.0017	0.1172	0.0105	0.1661	0.1518
216_QR_+7														0.0041	0.0113	0.0105	0.0297	0.0255
216_QR_+8															0.6391	0.7177	0.6406	0.8608
216_QR_+5																	0.4211	0.8283
216_Q_+1																	0.313	0.1105
216_Q_2																		0.5371
216_Q_1																		
216_U_1																		
216_L_+1																		
216_L_1																		
216_L_2																		
216_V_+2																		
216_V_+1																		
216_L_1																		
216_G_+1																		
216_F_+1																		
216_F_1																		
216_D_1																		
216_D_2																		
CNTL	67.80%	9.50%	94.70%	79.20%	76.60%	94.20%	93.90%	82.90%	35.30%	23.00%	89.60%	92.00%	92.60%	7.80%	75.70%	98.70%	50.00%	57.10%
CASE	46.40%	10.70%	100.00%	75.00%	75.00%	96.40%	96.40%	78.60%	32.10%	28.60%	96.40%	76.90%	78.60%	28.60%	71.40%	100.00%	39.30%	54.30%

TABLE 23 (CON'T)

	216_Q_+1	216_Q_2	216_Q_1	216_U_-1	216_L_+1	216_L_1	216_L_-1	216_L_2	216_V_+2	216_V_-1	216_L_1	216_G_-1	216_F_+1	216_F_1	216_D_1	216_D_2	
216_I_2	0.1309	0.1647	0.2214	0.1925	0.0531	0.0936	0.0568	0.0918	0.1241	0.1621	0.0908	0.1637	0.048	0.0615	0.0472	0.0683	216_I_2
216_I_3	0.4732	0.1073	0.88	0.9027	0.1186	0.9966	0.1189	0.9483	0.5519	0.8271	0.1788	0.9677	0.1256	0.8419	0.9949	0.8539	216_I_3
216_I_4	0.2716	0.8201	0.4625	0.4641	0.0218	0.3384	0.0215	0.4347	0.2913	0.8791	0.0425	0.3741	0.0476	0.9752	0.2527	0.3207	216_I_4
216_I_5	0.4941	0.1575	0.6092	0.1038	0.0531	0.8202	0.0576	0.1218	0.5548	0.1185	0.0652	0.2078	0.0448	0.1273	0.7435	0.6039	216_I_5
216_I_6	0.5361	0.0844	0.9136	0.8376	0.0579	0.9789	0.0598	0.9969	0.3487	0.835	0.0677	0.957	0.0943	0.8986	0.9784	0.7357	216_I_6
216_I_7	0.6372	0.037	0.8639	0.6497	0.0786	0.5075	0.0789	0.9077	0.6325	0.9561	0.1003	0.819	0.0757	0.831	0.5217	0.6292	216_I_7
216_I_8	0.6365	0.0261	0.3702	0.4536	0.0702	0.6191	0.075	0.8249	0.6225	0.7192	0.3281	0.7646	0.2019	0.8787	0.538	0.7074	216_I_8
216_I_+1	0.4903	0.046	0.319	0.5365	0.059	0.7057	0.0622	0.7784	0.8165	0.8129	0.1714	0.5849	0.0594	0.8532	0.7344	0.7752	216_I_+1
216_I_+2	0.3989	0.0897	0.8938	0.8403	0.069	0.7111	0.0671	0.7024	0.5255	0.8199	0.067	0.5879	0.1173	0.5041	0.2734	0.591	216_I_+2
216_I_+4	0.5765	0.0192	0.7132	0.6565	0.0325	0.6815	0.0345	0.7815	0.4893	0.4012	0.0964	0.1258	0.1161	0.1538	0.0672	0.1178	216_I_+4
216_R_+2	0.1157	0.0789	0.5406	0.4399	0.0636	0.3328	0.0661	0.111	0.486	0.1427	0.0515	0.1554	0.0901	0.1632	0.0676	0.1272	216_R_+2
216_R_+1	0.1518	0.0838	0.1592	0.1091	0.0633	0.1175	0.0602	0.179	0.486	0.1427	0.0515	0.1554	0.0901	0.1632	0.0676	0.1272	216_R_+1
216_R_2	0.1623	0.0866	0.1983	0.1344	0.0739	0.1472	0.0161	0.2042	0.0378	0.1459	0.1696	0.2001	0.0901	0.1632	0.0676	0.1272	216_R_2
216_R_1	0.0222	0.0112	0.0276	0.0203	0.0047	0.0159	0.0039	0.0213	0.0025	0.0152	0.0437	0.0176	0.0249	0.021	0.0054	0.0127	216_R_1
216_QR_+7	0.6188	0.0587	0.6882	0.8618	0.0368	0.7941	0.039	0.8881	0.7451	0.7462	0.253	0.9131	0.1308	0.6922	0.7005	0.6339	216_QR_+7
216_QR_+6	0.3013	0.0784	0.8185	0.6332	0.0591	0.4097	0.0702	0.9827	0.5193	0.7659	0.0859	0.9127	0.0677	0.766	0.9587	0.6737	216_QR_+6
216_QR_+5	0.2621	0.0591	0.4573	0.6112	0.1063	0.4489	0.106	0.589	0.4606	0.3573	0.1808	0.6586	0.1622	0.4067	0.2538	0.3683	216_QR_+5
216_QR_+4	0.1314	0.1379	0.7941	0.9027	0.0612	0.5826	0.0816	0.8155	0.3457	0.7442	0.0755	0.7063	0.2432	0.7513	0.5208	0.161	216_QR_+4
216_Q_+1	0.2724	0.0208	0.3082	0.4922	0.0923	0.289	0.092	0.465	0.4652	0.6406	0.1208	0.5445	0.1143	0.662	0.1418	0.2809	216_Q_+1
216_Q_2	0.0233	0.0233	0.1055	0.041	0.0587	0.0575	0.0558	0.0817	0.0825	0.0789	0.0831	0.1175	0.1098	0.0466	0.0377	0.0583	216_Q_2
216_Q_1	0.725	0.725	0.1276	0.3265	0.1276	0.794	0.1285	0.5428	0.8174	0.7175	0.3191	0.586	0.144	0.9111	0.7657	0.7839	216_Q_1
216_U_-1	0.591	0.591	0.0522	0.591	0.0522	0.6716	0.0532	0.7612	0.8174	0.7175	0.3191	0.586	0.144	0.9111	0.7657	0.7839	216_U_-1
216_L_+1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_L_+1
216_L_1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_L_1
216_L_2	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_L_2
216_V_+2	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_V_+2
216_V_+1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_V_+1
216_G_-1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_G_-1
216_F_+1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_F_+1
216_D_1	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_D_1
216_D_2	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	0.0143	216_D_2
CNTL	48.10%	75.30%	89.60%	83.10%	92.00%	0.60%	92.00%	6.70%	70.80%	94.20%	87.20%	8.20%	67.40%	94.80%	0.00%	0.70%	CNTL
CASE	62.50%	53.60%	87.50%	78.60%	75.00%	0.00%	75.00%	7.10%	60.70%	96.40%	71.40%	7.10%	46.40%	96.40%	0.00%	0.00%	CASE

**EXAMPLE 14: Transmission Disequilibrium Test (TDT)**

To ensure that the significant association observed in the case-control studies was not an artifact due to population admixture, a family based test of association, the transmission disequilibrium test (TDT) was conducted. By  
5 selecting a single affected offspring in each family, the TDT test performed a test of association (due to linkage disequilibrium) in the presence of linkage.

The test determined whether a particular allele was preferentially transmitted to an affected individual over what would be expected by chance. Only heterozygous parents were considered informative for the TDT. In addition, to  
10 increase power, heterozygous parents transmitting a different allele to two affected offspring were ignored. Accordingly, the TDT would be based on the same families that contributed to the linkage signal. The significance levels were estimated by Markov Chain Monte Carlo simulation methods as implemented in TDTEX from the S.A.G.E. program (Department of  
15 Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University, Cleveland, OH (1997)).

1. Asthma Phenotype: Five candidate SNPs were typed in the extended population in order to confirm the association seen in the case-  
20 control study. The five SNPs were in Gene 216 exons T5, T8, T+1, R1, and Q1. Since only heterozygote parents contribute information to the TDT test, SNP haplotypes (all 2-at-a-time and all 3-at-a-time) were constructed based on family data with the program GENEHUNTER (Kruglyak et al., 1996) in addition to analyzing the SNPs separately. This served to increase the informativeness  
25 of the single SNPs. These haplotypes were then used as "alleles" in future TDT analyses. In addition, p-values obtained from the TDT analyses were compared to the p-values obtained from the haplotyping in the case/control setting. To check for consistency, the p-values were recorded to compare the haplotype frequencies between the cases and controls of the over-transmitted  
30 alleles/haplotypes.

The TDT results strongly supported the association previously observed

in the case control studies (Table 24). Three of the five SNPs showed alleles that were preferentially transmitted to affected offspring ( $p < 0.04$  to  $< 0.0044$ ) in either the combined or UK population. When these SNPs were haplotyped together, most combinations had a haplotype that was preferentially transmitted to affected offspring ( $p < 0.03$  to  $< 0.001$ ). The most significant haplotype in the combined population was composed of SNPs T+1/R1/Q1 ( $p = 0.0006$ ). The most significant haplotype in the UK population was composed of SNPs T5/R1/Q1 ( $p = 0.0005$ ). In contrast to the UK population, none of the single SNP allele or multiple SNP haplotypes were preferentially over-transmitted to affected offspring at significant levels in the US population. This is most likely due to the combination of reduced power of the TDT versus the case-control study and the smaller sample size in the US.

Importantly, for all of the single SNP or multiple SNP haplotypes the allele that was significantly over-transmitted in either the combined population or in the UK sample was more frequent in the cases than in the controls. A summary of the TDT analyses and a comparison between the Case/control and TDT results are presented in Table 24.

2. Bronchial Hyper-responsiveness: The TDT analyses were repeated using only those asthmatic pairs that satisfied the additional criteria of having a  $PC_{20} \leq 16$  mg/ml (Table 25). The vast majority of single SNP and multiple SNP haplotypes showed increased significance with the more restricted phenotype. P values reached levels of  $< 0.00008$  for T5/R1/Q1 in the combined population and  $p < 0.000008$  in the UK sample. Similar to the yes/no phenotype, for the majority of the alleles in both the combined and UK population, the over-transmitted alleles in the TDT were more frequent in the cases. Similar to the yes/no phenotype with the less powerful TDT test, no significant results were observed with smaller US sample. In summary, the analysis of single SNPs and SNP haplotypes by the TDT test provided confirmatory evidence for Gene 216 as an asthma susceptibility gene.

**TABLE 24**

Asthma Yes/NO				
Combined US and UK				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.0337	0.0213	89.5%	94.8%
R 1	0.0725	NS	88.7%	88.2%
T +1	0.0956	0.0055	85.2%	92.4%
T 8	1.0000	NS	NA	
T 5	0.1364	0.0420	76.7%	83.3%
R1Q1	0.0042	0.1362	78.2%	83.1%
T+1Q1	0.0932	0.0049	85.2%	92.4%
T8Q1	0.0553	0.0084	86.0%	92.9%
T5Q1	0.2659	0.0342	76.2%	83.0%
T+1R1	0.0029	0.0465	73.9%	80.6%
T8R1	0.0799	NS	85.1%	67.9%
T5R1	0.0107	0.1537	66.1%	71.5%
T8T+1	0.2762	0.0044	85.2%	92.4%
T5T+1	0.3078	0.0012	72.5%	83.0%
T5T8	0.0948	0.0028	73.7%	83.4%
T+1R1Q1	0.0006	0.0430	73.9%	80.8%
T8R1Q1	0.0086	0.0552	74.7%	81.2%
T5R1Q1	0.0025	0.1591	65.9%	71.2%
T5T+1R1	0.0136	0.0175	62.3%	71.2%
T8T+1R1	0.0084	0.0377	73.9%	80.9%
T5T8R1	0.0060	0.0235	63.0%	71.5%
T5T8Q1	0.1242	0.0033	73.1%	83.0%
T5T8T+1	0.1540	0.0009	72.7%	83.0%
T8T+1Q1	0.1351	0.0043	85.3%	92.4%
T5T+1Q1	0.1080	0.0010	72.5%	83.0%

NS = non-significant or over-transmitted allele not present more often in cases than controls  
 NA = no alleles were over-transmitted

**TABLE 24 (CON'T)**

Asthma Yes/NO				
UK				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.0044	0.0274	89.4%	95.1%
R 1	0.3665	0.1473	86.8%	91.4%
T +1	0.0128	0.0105	86.4%	93.8%
T 8	1.0000	NS	NA	
T 5	0.0434	0.0426	75.4%	83.3%
R1Q1	0.0044	0.0069	76.2%	86.5%
T+1Q1	0.0714	0.0066	86.4%	93.8%
T8Q1	0.0342	0.0275	87.4%	93.6%
T5Q1	0.1687	0.0314	74.9%	82.9%
T+1R1	0.0269	0.0018	73.2%	85.1%
T8R1	0.4848	0.0933	84.6%	89.9%
T5R1	0.0639	0.0067	63.1%	74.7%
T8T+1	0.2254	0.0069	86.4%	93.8%
T5T+1	0.2007	0.0088	72.9%	82.9%
T5T8	0.0277	0.0103	73.7%	83.4%
T+1R1Q1	0.0063	0.0016	73.2%	85.1%
T8R1Q1	0.0139	0.0039	74.1%	85.0%
T5R1Q1	0.0005	0.0136	63.4%	74.2%
T5T+1R1	0.0220	0.0036	61.5%	74.2%
T8T+1R1	0.0043	0.0012	73.2%	85.1%
T5T8R1	0.0095	0.0018	61.5%	74.7%
T5T8Q1	0.0074	0.0105	73.3%	82.9%
T5T8T+1	0.0255	0.0082	73.0%	82.9%
T8T+1Q1	0.0207	0.0087	86.4%	93.8%
T5T+1Q1	0.0127	0.0093	72.9%	82.9%

NS = non-significant or over-transmitted allele not present more often in cases than controls;  
 NA = no alleles were over-transmitted

**TABLE 24 (CON'T)**

Asthma Yes/NO				
US				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.8039	NS	10.4%	6.3%
R 1	0.1067	NS	92.2%	75.9%
T +1	0.6288	NS	17.1%	13.0%
T 8	1.0000	NS	NA	
T 5	0.7020	NS	20.8%	16.7%
R1Q1	0.2134	NS	81.8%	69.6%
T+1Q1	0.6811	NS	10.4%	9.7%
T8Q1	0.7584	0.2887	83.6%	90.2%
T5Q1	0.8284	NS	9.7%	8.3%
T+1R1	0.0658	NS	75.1%	63.0%
T8R1	0.0687	NS	86.1%	72.2%
T5R1	0.1859	NS	71.4%	59.3%
T8T+1	0.9465	0.4778	83.0%	87.0%
T5T+1	0.8537	0.5074	9.7%	13.0%
T5T8	0.8848	NS	20.8%	13.0%
T+1R1Q1	0.1569	NS	75.2%	62.7%
T8R1Q1	0.2386	NS	75.8%	66.0%
T5R1Q1	0.0831	NS	70.7%	59.3%
T5T+1R1	0.1332	NS	64.1%	59.9%
T8T+1R1	0.1299	NS	75.2%	63.4%
T5T8R1	0.0813	NS	65.5%	60.2%
T5T8Q1	0.8654	NS	9.7%	7.8%
T5T8T+1	0.8546	NS	9.6%	9.3%
T8T+1Q1	0.6864	NS	10.4%	9.3%
T5T+1Q1	0.8618	0.9991	9.7%	9.7%

NS = non-significant or over-transmitted allele not present more often in cases than controls;  
 NA = no alleles were over-transmitted

**TABLE 25**

BHR				
Combined US and UK				
Over-Transmitted Haplotype				
Exon in Gene 216	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
Q 1	0.0800	0.1565	89.5%	94.2%
R 1	0.0374	NS	88.7%	88.3%
T +1	0.1252	0.1413	85.2%	90.6%
T 8	1.0000	NS	NA	
T 5	0.0947	0.4681	76.7%	80.2%
R1Q1	0.0017	0.2040	78.2%	83.7%
T+1Q1	0.1835	0.1192	85.2%	90.6%
T8Q1	0.1616	0.0987	86.0%	91.8%
T5Q1	0.1496	0.3214	76.2%	80.2%
T+1R1	0.0015	0.1479	73.9%	80.2%
T8R1	0.0281	0.7994	85.1%	85.9%
T5R1	0.0009	0.6419	66.1%	68.4%
T8T+1	0.6224	0.1380	85.2%	90.6%
T5T+1	0.4821	0.0660	72.5%	80.3%
T5T8	0.1786	0.1284	73.7%	80.2%
T+1R1Q1	0.0003	0.1426	73.9%	80.4%
T8R1Q1	0.0035	0.1298	74.7%	81.4%
T5R1Q1	0.0001	0.4524	65.9%	69.7%
T5T+1R1	0.0052	0.1332	62.3%	69.6%
T8T+1R1	0.0066	0.1397	73.9%	80.6%
T5T8R1	0.0028	0.2632	63.0%	68.4%
T5T8Q1	0.3680	0.0954	73.1%	80.3%
T5T8T+1	0.5282	0.0786	72.7%	80.3%
T8T+1Q1	0.3105	0.1261	85.3%	90.6%
T5T+1Q1	0.5276	0.0686	72.5%	80.3%

NS = non-significant or over-transmitted allele not present more often in cases than controls;  
 NA = no alleles were over-transmitted



**TABLE 25 (CON'T)**

BHR				
UK				
Over-Transmitted Haplotype				
Exon in Gene	TDT p-value	Case/Control p-value	Control Frequency	Case Frequency
216				
Q 1	0.0069	0.0613	89.4%	95.8%
R 1	0.3285	0.1041	86.8%	93.0%
T +1	0.0201	0.0454	86.4%	94.0%
T 8	1.0000	NS	NA	
T 5	0.0367	0.2644	75.4%	81.6%
R1Q1	0.00078	0.0052	76.2%	89.8%
T+1Q1	0.0209	0.0280	86.4%	94.0%
T8Q1	0.0120	0.0933	87.4%	93.8%
T5Q1	0.0974	0.1624	74.9%	81.7%
T+1R1	0.0001	0.0026	73.2%	87.6%
T8R1	0.2818	0.1182	84.6%	91.0%
T5R1	0.0038	0.0420	63.1%	74.6%
T8T+1	0.1437	0.0327	86.4%	94.0%
T5T+1	0.0902	0.0739	72.9%	81.7%
T5T8	0.0536	0.1052	73.7%	81.7%
T+1R1Q1	0.000075	0.0042	73.2%	87.8%
T8R1Q1	0.0031	0.0056	74.1%	87.7%
T5R1Q1	0.0000078	0.0331	63.4%	75.4%
T5T+1R1	0.0071	0.0131	61.5%	75.3%
T8T+1R1	0.0023	0.0034	73.2%	87.8%
T5T8R1	0.0073	0.0216	61.5%	74.6%
T5T8Q1	0.0424	0.0835	73.3%	81.7%
T5T8T+1	0.1380	0.0761	73.0%	81.7%
T8T+1Q1	0.0322	0.0319	86.4%	94.0%
T5T+1Q1	0.1096	0.0756	72.9%	81.7%

NS = non-significant or over-transmitted allele not present more often in cases than controls;  
 NA = no alleles were over-transmitted

**TABLE 25 (CON'T)**

BHR

US

Exon in Gene 216	TDT p-value	Over-Transmitted Haplotype		
		Case/Control p- value	Control Frequency	Case Frequency
Q_1	0.5081	0.7250	10.4%	12.5%
R_1	0.0577	NS.	92.2%	71.4%
T_+1	0.5493	0.5937	17.1%	21.4%
T_8	1.0000	NS	NA	
T_5	0.7741	0.6206	20.8%	25.0%
R1Q1	0.1259	NS	81.8%	58.8%
T+1Q1	0.7495	0.1224	10.4%	21.4%
T8Q1	0.7514	0.7864	10.4%	12.1%
T5Q1	0.1029	0.1408	9.7%	18.8%
T+1R1	0.2012	NS	75.1%	50.0%
T8R1	0.0880	NS	86.1%	67.9%
T5R1	0.0963	NS	71.4%	46.4%
T8T+1	0.7557	0.2626	10.7%	17.9%
T5T+1	0.4904	0.0908	9.7%	21.4%
T5T8	0.8871	0.9876	20.8%	21.4%
T+1R1Q1	0.0828	NS	75.2%	50.0%
T8R1Q1	0.1759	NS	75.8%	55.9%
T5R1Q1	0.2046	NS	70.7%	46.4%
T5T+1R1	0.1915	NS	64.1%	46.4%
T8T+1R1	0.2537	NS	75.2%	50.0%
T5T8R1	0.1633	NS	65.5%	46.4%
T5T8Q1	0.6920	0.3863	9.7%	16.1%
T5T8T+1	0.8586	0.3158	9.6%	17.9%
T8T+1Q1	0.7517	0.3367	10.4%	17.9%
T5T+1Q1	0.8579	0.1166	9.7%	21.4%

NS = non-significant or over-transmitted allele not present more often in cases than controls;

NA = no alleles were over-transmitted

**EXAMPLE 15: Attributable Risk Assessment**

From the knowledge of the frequency of a functional polymorphism and the relative risk of the heterozygote and homozygote (at-risk) genotypes, one can evaluate the attributable fraction (M.J. Khoury et al., 1993, Fundamentals of Genetic Epidemiology, J.L. Kelsey et al., (eds), *Monographs in Epidemiology and Biostatistics*, Oxford University Press, New York, NY, Section 3, pp 74-77) or attributable risk in the population. An attributable fraction of 25% would mean that if the population were monomorphic for the protective allele, the prevalence of the trait would be 25% lower.

The formula for the attributable fraction is:

$$\text{Attributable fraction} = \frac{(1-f)^2 + 2f(1-f)\gamma + f^2\eta - 1}{(1-f)^2 + 2f(1-f)\gamma + f^2\eta},$$

where  $f$  is the allele frequency,  $\gamma$  is the relative risk of the heterozygote genotype over the wild type homozygote, and  $\eta$  is the risk of the homozygote mutant over the wild type homozygote. This approach requires the estimation of  $f$ ,  $\gamma$  and  $\eta$ . Ideally these quantities should be estimated in an epidemiological sample.

The study design (genome scan with affected sibling pairs followed by association study using IBD = 2 individuals as cases in the case/control comparison) offers maximum power to detect linkage and association, but does not provide estimates of the required parameters, namely 1) the relative risk (or odds ratio) of the genotype/allele for most SNPs or haplotypes and 2) the frequency of the SNP in the general population. In a recent paper, Altshuler et al. used the data from a TDT analysis to estimate allele and genotype relative risks assuming a multiplicative model or  $\eta = \gamma^2$  (D. Altshuler et al., 2000, *Nature Genetics* **26**:76-80). Thus, the mutant homozygote is predicted to carry a relative risk equal to the square of the risk for the heterozygote.

To overcome some of the difficulties mentioned above that are associated with a case/control design, the data obtained from typing 5 SNPs in Gene 216 on the entire population (not just the subset of IBD = 2 individuals)

were used to estimate the relative risk of these 5 SNPs. The data from the TDT obtained by using the first asthmatic sibling per family were used. Because of the limited number of informative matings in the TDT analysis, a multiplicative model for the genotype relative risk was used as in the Altshuler et. al paper, i.e.  $\eta = \gamma^2$ . An interval on the attributable fraction estimates was made by constructing individual confidence regions for the allele frequency in the control population and for the attributable risk obtained from the TDT data. While combining these two confidence intervals to obtain a confidence region for the attributable fraction did not lead to a proper confidence region with the required coverage, it determined the variability involved in estimating the attributable fraction. As a short hand notation, this is referred to as a confidence interval with coverage equal to the one used for the constituent parameters.

By using the control population to estimate allele frequencies, the attributable risk was underestimated. Based on these assumptions, the attributable risk for the single SNPs that were significant in the case-control study ( $p < 0.05$ ) in either population was computed. The AF was also computed for all SNP combinations significant in the combined TDT analysis ( $p < 0.01$ ) using the asthma phenotype. These values are shown below.

SNP(s)	Attributable fraction (AF) estimate	80% Confidence Interval
Q 1	50%	17 to 65%
R 1	37%	4 to 57%
T + 1	39%	7 to 57%
T 5	22%	0 to 35%
R1 Q1	36%	14 to 54%
T +1 R1	29%	8 to 47%
T +1 R1 Q1	34%	14 to 52%
T 5 R1 Q1	19%	3 to 38%
T 5 T 8 R1	24%	9 to 41%
T 8 R1 Q1	32%	11 to 50%
T 8 T+1 R1	25%	2 to 44%

Because the alleles that confer increased risk of developing asthma are so common (haplotype frequencies ranging from 60% to 83%), their effect translated into a substantial population attributable risk, with estimates ranging from 19 to 50% for different SNPs or SNP haplotypes. These computations  
5 depended heavily on allele frequency and risk estimates. Proper estimates of the attributable fraction are based on a population sample and are only meaningful for functional SNPs or SNP haplotypes.

Conclusion: Gene 216 has been demonstrated to be an asthma gene in accordance with the data disclosed herein, including: 1) localization to a  
10 region on chromosome 20 identified through linkage; 2) polymorphism analysis performed to identify sequence variants localized in the candidate gene; 3) genotype analyses of the identified polymorphisms; 4) association between identified alleles and the asthma phenotype in a case-control analysis; 5) association between identified alleles and the asthma phenotype in  
15 transmission disequilibrium tests (TDT), haplotype analyses, and analyses using additional phenotypes; 6) identification of transcripts in tissues relevant to pulmonary disease and/or inflammation; and 7) characterization of Gene 216 as an ADAM family member. In addition to respiratory diseases, Gene 216 is likely to be involved in obesity and inflammatory bowel disease, as  
20 obesity (Wilson et al., 1999, *Arch. Intern. Med.* **159**: 2513-14) and inflammatory bowel disease (B. Wallaert et al., 1995, *J. Exp. Med.* **182**:1897-1904) have been linked to asthma.

#### **EXAMPLE 16: Protein Expression And Purification**

Expression and purification of the Gene 216 protein of the invention can  
25 be performed essentially as outlined below. To facilitate the cloning, expression, and purification of membrane and secreted protein from the 20p13-p12, a gene expression system, such as the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli* is selected. Also, a DNA sequence encoding a peptide tag, the His-Tap, is fused to the 3' end of  
30 DNA sequences of interest to facilitate purification of the recombinant protein products. The 3' end is selected for fusion to avoid alteration of any 5' terminal

signal sequence.

Nucleic acids chosen, for example, from the nucleic acids set forth in SEQ ID NO:1 or SEQ ID NO:6 (Figures 24 and 29, respectively) for cloning the genes are prepared by polymerase chain reaction (PCR). Synthetic oligonucleotide primers specific for the 5' and 3' ends of the nucleotide sequences are designed and purchased from Life Technologies. All forward primers (specific for the 5' end of the sequence) are designed to include an *Nco*I cloning site at the 5' terminus. These primers are designed to permit initiation of protein translation at the methionine residue encoded within the *Nco*I site followed by a valine residue and the protein encoded by the DNA sequence. All reverse primers (specific for the 3' end of the sequence) include an *Eco*RI site at the 5' terminus to permit cloning of the sequence into the reading frame of the pET-28b. The pET-28b vector provides a sequence encoding an additional 20 carboxyl-terminal amino acids including six histidine residues (at the C-terminus), which comprise the histidine affinity tag.

DNA prepared from the 20p13-p12 region is used as the source of template DNA for PCR amplification (Ausubel et al., 1994). To amplify a DNA sequence containing the nucleotide sequence, c DNA (50 ng) is introduced into a reaction vial containing 2 mM MgCl<sub>2</sub>, 1 μM synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined 20p13-p12 region, 0.2 mM of each of deoxynucleotide triphosphate, dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ) in a final volume of 100 microliters.

Upon completion of thermal cycling reactions, each sample of amplified DNA is purified using the Qiaquick Spin PCR purification kit. All amplified DNA samples are subjected to digestion with the restriction endonucleases, e.g., *Nco*I and *Eco*RI (NEB) (Ausubel et al., 1994). DNA samples are then subjected to electrophoresis on 1.0% NuSeive (FMC BioProducts) agarose gels. DNA is visualized by exposure to ethidium bromide and long wave UV irradiation. DNA contained in slices isolated from the agarose gel was purified using the BIO 101 GeneClean Kit protocol.

The pET-28b vector is prepared for cloning by digestion with restriction endonucleases, e.g., *NcoI* and *EcoRI* (NEB) (Ausubel et al., 1994). The pET-28a vector, which encodes the histidine affinity tag that can be fused to the 5' end of an inserted gene, is prepared by digestion with appropriate restriction  
5 endonucleases.

Following digestion, DNA inserts are cloned (Ausubel et al., 1994) into the previously digested pET-28b expression vector. Products of the ligation reaction are then used to transform the BL21 strain of *E. coli* (Ausubel et al., 1994) as described below.

10 Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21 (DE3), are transformed with recombinant pET expression plasmids carrying the cloned sequence according to standard methods (Ausubel et al., 1994). Briefly, 1 microliter of ligation reaction is mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which samples were  
15 incubated in 0.45 ml SOC medium (0.5% yeast extract, 2.0% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub> and 20 mM glucose) at 37°C with shaking for 1 hr. Samples are then spread on LB agar plates containing 25 µg/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 are then picked and analyzed to evaluate cloned inserts, as described  
20 below.

Individual BL21 clones transformed with recombinant pET-28b. 20p13-p12 region nucleotide sequences are analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers specific for the 20p13-p12 region sequences that are used in the original PCR amplification  
25 cloning reactions. Successful amplification verifies the integration of the sequence in the expression vector (Ausubel et al., 1994).

Individual clones of recombinant pET-28b vectors carrying properly cloned 20p13-p12 region nucleotide sequences are picked and incubated in 5 ml of LB broth plus 25 µg/ml kanamycin sulfate overnight. The following day  
30 plasmid DNA is isolated and purified using the QIAGEN plasmid purification protocol.

The pET vector can be propagated in any *E. coli* K-12 strain, e.g., HMS174, HB101, JM109, DH5, and the like, for purposes of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are  
5 lysogens of bacteriophage DE3, a lambda derivative that carries the *lacI* gene, the *lacUV5* promoter, and the gene for T7 RNA polymerase. T7 RNA polymerase is induced by addition of isopropyl- $\beta$ -D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid containing a functional T7 promoter, such as pET-28b, carrying its gene of interest. Strains include,  
10 for example, BL21(DE3) (Studier et al., 1990, *Meth. Enzymol.*, **185**:60-89).

To express the recombinant sequence, 50 ng of plasmid DNA are isolated as described above to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression kit). The *lacZ* gene ( $\beta$ -galactosidase) is expressed in the pET-System as described for  
15 the 20p13-p12 region recombinant constructions. Transformed cells were cultured in SOC medium for 1 hr, and the culture is then plated on LB plates containing 25  $\mu$ g/ml kanamycin sulfate. The following day, the bacterial colonies are pooled and grown in LB medium containing kanamycin sulfate (25  $\mu$ g/ml) to an optical density at 600 nm of 0.5 to 1.0 OD units, at which point 1  
20 mM IPTG was added to the culture for 3 hr to induce gene expression of the 20p13-p12 region recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria are collected by centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 min at 4°C. Pellets are resuspended in 50 ml of cold mM Tris-HCl, pH 8.0, 0.1 M NaCl and  
25 0.1 mM EDTA (STE buffer). Cells are then centrifuged at 2000 x g for 20 min at 4°C. Wet pellets are weighed and frozen at -80°C until ready for protein purification.

A variety of methodologies known in the art can be used to purify the isolated proteins (Coligan et al., 1995, *Current Protocols in Protein Science*,  
30 John Wiley & Sons, New York, NY). For example, the frozen cells can be thawed, resuspended in buffer, and ruptured by several passages through a



small volume microfluidizer (Model M-110S, Microfluidics International Corp., Newton, MA). The resultant homogenate is centrifuged to yield a clear supernatant (crude extract) and, following filtration, the crude extract is fractioned over columns. Fractions are monitored by absorbance at OD<sub>280</sub> nm and peak fractions may be analyzed by SDS-PAGE.

The concentrations of purified protein preparations are quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, 1986, *Eur. J. Biochem.*, **157**:169-180). Protein concentrations are also measured by the method of Bradford, 1976, *Anal. Biochem.*, **72**:248-254; and Lowry et al., 1951, *J. Biol. Chem.*, **193**:265-275 using bovine serum albumin as a standard.

SDS-polyacrylamide gels of various concentrations are purchased from Bio-Rad, and stained with Coomassie blue. Molecular weight markers may include rabbit skeletal muscle myosin (200 kDa), *E. coli*  $\beta$ -galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

Proteins can also be isolated by other conventional means of protein biochemistry and purification to obtain a substantially pure product, i.e., 80, 95, or 99% free of cell component contaminants, as described in Jacoby, 1984, *Methods in Enzymology*, Vol. 104, Academic Press, NY; Scoopes, 1987, *Protein Purification, Principles and Practice*, 2<sup>nd</sup> Ed., Springer-Verlag, NY; and Deutscher (ed), 1990, *Guide to Protein Purification, Methods in Enzymology*, Vol. 182. If the protein is secreted, it can be isolated from the supernatant in which the host cell is grown; otherwise, it can be isolated from a lysate of the host cells.

Once a sufficient quantity of the desired protein has been obtained, it may be used for various purposes. One use of the protein or polypeptide is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo*

techniques well known in the art. Monoclonal antibodies to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas (Kohler, 1975, *Nature*, **256**:495). In summary, a mouse is inoculated with a few micrograms of protein over a period of 2 weeks. The mouse is then sacrificed. The cells that produce antibodies are then removed from the mouse's spleen. The spleen cells are then fused with polyethylene glycol with mouse myeloma cells. The successfully fused cells are diluted in a microtiter plate and growth of the culture is continued. The amount of antibody per well is measured by immunoassay methods such as ELISA (Engvall, 1980, *Meth. Enzymol.*, **70**:419). Clones producing antibody can be expanded and further propagated to produce protein antibodies. Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors. See Huse et al., 1989, *Science*, **246**:1275-1281. For additional information on antibody production see Davis et al., 1989, *Basic Methods in Molecular Biology*, Elsevier, NY, Section 21-2. Such antibodies are particularly useful in diagnostic assays for detection of variant protein forms, or as an active ingredient in a pharmaceutical composition.

The disclosure of each of the patents, patent applications, and publications cited in the specification is hereby incorporated by reference herein in its entirety.

Although the invention has been set forth in detail, one skilled in the art will recognize that numerous changes and modifications can be made, and that such changes and modifications may be made without departing from the spirit and scope of the invention.

**WHAT IS CLAIMED IS:**

1. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:
  - 5 a. SEQ ID NO:1;
  - b. a nucleotide sequence encoding amino acid SEQ ID NO:4;
  - c. a nucleotide sequence complementary to SEQ ID NO:1;
  - d. a nucleotide sequence which hybridizes under high stringency conditions to SEQ ID NO:1;
  - 10 e. a nucleotide sequence which hybridizes under moderate stringency conditions to SEQ ID NO:1;
  - f. a nucleotide sequence which hybridizes under low stringency conditions to SEQ ID NO:1;
  - g. a nucleotide sequence which is at least 95% identical to the
  - 15 sequence of SEQ ID NO:1;
  - h. a nucleotide sequence which is at least 80% identical to the sequence of SEQ ID NO:1; and
  - i. a nucleotide sequence which is at least 50% identical to the sequence of SEQ ID NO:1.
- 20 2. The isolated nucleic acid of claim 1 which is DNA.
3. The isolated nucleic acid of claim 1 which is RNA.
4. A vector comprising the isolated nucleic acid of claim 1.
5. A host cell comprising the expression vector of claim 4.
6. The host cell of claim 5 which is selected from the group
- 25 consisting of eukaryotic and prokaryotic cells.
7. The host cell of claim 5 which is selected from the group consisting of bacterial, fungal.
8. The isolated nucleic acid of claim 1, wherein the nucleic acid

sequence comprises at least 50 consecutive nucleotides.

9. A vector comprising the isolated nucleic acid of claim 8.
10. A host cell comprising the expression vector of claim 9.
11. The host cell of claim 10 which is selected from the group  
5 consisting of eukaryotic and prokaryotic cells.
12. The host cell of claim 10 which is selected from the group consisting of bacterial, yeast, insect, mammalian, and plant cells.
13. The isolated nucleic acid of claim 1, wherein the nucleic acid sequence comprises at least 15 consecutive nucleotides.
- 10 14. A vector comprising the isolated nucleic acid of claim 13.
15. A host cell comprising the vector of claim 14.
16. The host cell of claim 15 which is selected from the group consisting of eukaryotic and prokaryotic cells.
17. The host cell of claim 15 which is selected from the group  
15 consisting of bacterial, yeast, insect, mammalian, and plant cells.
18. An isolated nucleic acid variant which comprises the sequence of SEQ ID NO:6, and contains at least one single nucleotide polymorphism set forth in Table 10.
19. An isolated nucleic acid variant which comprises at least 50  
20 consecutive nucleotides of SEQ ID NO:6, and contains at least one single nucleotide polymorphism set forth in Table 10.
20. An isolated nucleic acid variant which comprises at least 15 consecutive nucleotides of SEQ ID NO:6, and contains at least one single nucleotide polymorphism set forth in Table 10.
- 25 21. The isolated nucleic acid variant of claim 20, wherein the single nucleotide polymorphism is selected from the group consisting of T4, T5, T8,

T+1, T+2, R1, Q1, Q2, QR+4, QR+6, QR+7, and U-1.

22. The isolated nucleic acid variant of claim 20, wherein the single nucleotide polymorphism selected from the group consisting of D1, F1, I1, L1, R2, T6, T1, T2, T3, and T7.

5 23. The isolated nucleic acid variant of claim 20 containing at least two single nucleotide polymorphisms selected from the group consisting of:

- a. T+2 and QR+4;
- b. QR+5 and QR+4;
- c. QR+4 and Q+1;
- 10 d. QR+6 and Q2; and
- e. QR+4 and Q2.

24. The isolated nucleic acid variant of claim 20, wherein the single nucleotide polymorphism is selected from the group consisting of:

- a. T5 and T8;
- 15 b. T+2 and QR+4;
- c. T4 and T5.
- d. T+1 and R1 and Q1; and
- e. T5 and R1 and Q1.

25 25. An isolated nucleic acid variant which comprises the sequence of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

26. An isolated nucleic acid variant which comprises at least 50 consecutive nucleotides of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

25 27. An isolated nucleic acid variant which comprises at least 15 consecutive nucleotides of SEQ ID NO:1, and contains at least one single nucleotide polymorphism at a site shown in Figure 24.

28. An isolated alternate splice variant which comprises at least one

exon of SEQ ID NO:1 set forth in Figures 9 and 10.

29. An isolated alternate splice variant which comprises at least one exon of SEQ ID NO:1 selected from the group consisting of exons T, R, Q, and U set forth in Figures 9 and 10.

5           30. An isolated alternate splice variant which comprises at least one exon of SEQ ID NO:1 selected from the group consisting of exons A, B, C, D, D', E, F, G, H, I, J, K, L, L2, M, N, O, P, and S set forth in Figures 9 and 10.

31. An isolated alternate splice variant which comprises a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:350-362.

10           32. An isolated polypeptide encoded by the nucleic acid of any one of claims 1 and 8.

33. An isolated polypeptide encoded by the nucleic acid of any one of claims 18, 19, 25, and 26.

15           34. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- a. SEQ ID NO:4;
- b. an amino acid sequence which is at least 80% identical to SEQ ID NO:4;
- c. an amino acid sequence which is at least 75% identical to SEQ ID NO:4; and
- 20           d. an amino acid sequence which is at least 65% identical to SEQ ID NO:4.

35. An isolated polypeptide comprising at least 20 consecutive residues of the amino acid sequence of claim 34.

25           36. An isolated polypeptide comprising at least 7 consecutive residues of the amino acid sequence of claim 34.

37. An antibody or antibody fragment which binds to the isolated

polypeptide of claim 32.

38. An antibody or antibody fragment which binds to the isolated polypeptide of claim 33.

39. An antibody or antibody fragment which binds to the isolated  
5 polypeptide according to any one of claims 34-36

40. The antibody or antibody fragment of claim 37 which is selected from the group consisting of polyclonal and monoclonal antibodies.

41. The antibody or antibody fragment of claim 38 which is selected from the group consisting of polyclonal and monoclonal antibodies.

10 42. The antibody or antibody fragment of claim 39 which is selected from the group consisting of polyclonal and monoclonal antibodies.

43. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- a. SEQ ID NO:6;
- 15 b. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:6; and
- c. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:6.

20 44. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- a. SEQ ID NO:364;
- b. a nucleotide sequence complementary to SEQ ID NO:364.
- c. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:364;
- 25 d. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:364.
- e. SEQ ID NO:365;
- f. a nucleotide sequence complementary to SEQ ID NO:365;

- g. a nucleotide sequence comprising at least 50 consecutive nucleotides of SEQ ID NO:365; and
  - h. a nucleotide sequence comprising at least 15 consecutive nucleotides of SEQ ID NO:365.
- 5           45. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
- a. SEQ ID NO:366;
  - b. an amino acid sequence comprising 20 consecutive residues of SEQ ID NO:366; and
  - 10          c. an amino acid sequence comprising 7 consecutive residues of SEQ ID NO:M366.
46. An isolated antibody or antibody fragment that binds to the isolated polypeptide of claim 45.
47. The antibody or antibody fragment of claim 46 which is selected  
15   from the group consisting of monoclonal and polyclonal antibodies.
48. An isolated antisense nucleic acid comprising at least 15 consecutive nucleotides of a sequence complementary to SEQ ID NO:1.
49. An isolated antisense nucleic acid comprising at least 15 consecutive nucleotides of a sequence complementary to SEQ ID NO:6.
- 20          50. A vector comprising the isolated antisense nucleic acid of any one of claims 48-49.
51. A kit for detecting a Gene 216 nucleotide sequence comprising:  
a. the isolated nucleic acid of any one of claims 13, 20, and 27; and  
b. at least one component to detect binding of the isolated nucleic  
25   acid to a Gene 216 nucleotide sequence.
52. A kit for detecting a Gene 216 amino acid sequence comprising:  
a. the isolated antibody of claim 42; and  
b. at least one component to detect binding of the isolated antibody



to a Gene 216 amino acid sequence.

53. A method of identifying a Gene 216 ligand, comprising:

- a. contacting the isolated polypeptide of claim 35 with a test agent under conditions that allow the polypeptide to bind to the test agent, and  
5 thereby form a complex; and  
b. detecting the polypeptide-test agent complex of (a), wherein detection of the complex indicates identification of a Gene 216 ligand.

54. The method of claim 53, wherein the ligand is a metalloprotease inhibitor.

10 55. The method of claim 54, wherein the metalloprotease inhibitor is a proglutamyl peptide analog.

56. The method of claim 55, wherein the proglutamyl peptide analog is an analog of pyroGlu-Asn-Trp-OH or pyroGlu-Glu-Trp-OH.

15 57. A pharmaceutical composition comprising the ligand identified according to the method of any one of claims 53-56, and a physiologically acceptable carrier, excipient, or diluent.

58. A pharmaceutical composition comprising the isolated nucleic acid of any one of claims 1, 8, 13, 43, 48, and 49, and a physiologically acceptable carrier, excipient, or diluent.

20 59. A pharmaceutical composition comprising the vector of any one of claims 4, 9, 14, and 48, and a physiologically acceptable carrier, excipient, or diluent.

60. A pharmaceutical composition comprising the isolated antibody or antibody fragment of claim 42, and a physiologically acceptable carrier,  
25 excipient, or diluent.

61. A pharmaceutical composition comprising the isolated polypeptide of claim 36 and a physiologically acceptable carrier, excipient, or

diluent.

62. A method of identifying a human Gene 216 or ortholog, comprising:

- 5 a. contacting the nucleic acid of any one of claims 1, 8, and 13 with a biological sample under conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex; and
- b. detecting the hybridization complex of (a), wherein detection of the complex indicates identification of a human Gene 216 or ortholog.

63. A method of treating a chromosome 20 disorder comprising  
10 administering the pharmaceutical composition of claim 57 in an amount effective to treat the disorder.

64. The method of claim 63, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

15 65. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 58 in an amount effective to treat the disorder.

66. The method of claim 65, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory  
20 bowel disease.

67. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 59 in an amount effective to treat the disorder.

68. The method of claim 67, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory  
25 bowel disease.

69. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 60 in an amount

effective to treat the disorder.

70. The method of claim 69, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

5        71. A method of treating a chromosome 20 disorder comprising administering the pharmaceutical composition of claim 61 in an amount effective to treat the disorder.

72. The method of claim 71, wherein the chromosome 20 disorder is selected from the group consisting of asthma, obesity, and inflammatory  
10        bowel disease.

73. A transgenic mouse whose genome comprises an introduced null mutation in an endogenous Gene 216.

74. The transgenic mouse of claim 73, wherein both alleles of the endogenous Gene 216 of said mouse have been disrupted.

15        75. The transgenic mouse of claim 74, wherein the mouse genome further comprises a human Gene 216 nucleic acid sequence.

76. A method of making a homozygous transgenic knockout mouse comprising:

- 20        a. disrupting an endogenous Gene 216 in mouse embryonic stem cells;
- b. introducing said embryonic stem cells into a mouse blastocyst and transplanting said blastocyst into a pseudopregnant mouse;
- c. allowing said blastocyst to develop into a chimeric mouse;
- d. breeding said chimeric mouse to produce offspring; and
- 25        e. screening said offspring to identify a homozygous transgenic knockout mouse.

77. A method of making a knockout mouse comprising administering the antibody or antibody fragment of claim 47 in an amount effective to disrupt

endogenous Gene 216 polypeptide function, thereby making a knockout mouse.

78. A method of forming a crystal of the isolated Gene 216 polypeptide of claim 36 comprising:

- 5           a. incubating the polypeptide with a solution selected from the group consisting of the solutions in wells 1-30 in Table 1 under conditions to allow crystalization; and
- b. detecting the crystalization in (a), whereby crystalization indicates formation of a Gene 216 polypeptide crystal.

10           79. A method of diagnosing a chromosome 20 disorder, comprising:

- a. contacting the isolated nucleic acid of any one of claims 20-24 with a biological sample under high stringency conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex; and
- 15           b. detecting the hybridization complex of (a), wherein detection of the complex indicates diagnosis of a chromosome disorder.

80. The method of claim 79, wherein the disorder is selected from the group consisting of asthma, obesity, and inflammatory bowel disease.

81. A method of diagnosing a chromosome 20 disorder comprising:

- 20           a. contacting the isolated antibody or antibody fragment of claim 41 with a biological sample under high stringency conditions that allow the antibody or antibody fragment to bind to an amino acid sequence in the sample, and thereby form a complex; and
- b. detecting the complex of (a), wherein detection of the complex
- 25           indicates diagnosis of a chromosome disorder.

82. A method of determining a pharmacogenetic profile comprising:

- a. contacting the isolated nucleic acid of any one of claims 20-24 with a biological sample under high stringency conditions that allow the nucleic acid to hybridize to a nucleic acid in the sample, and thereby form a complex;

and

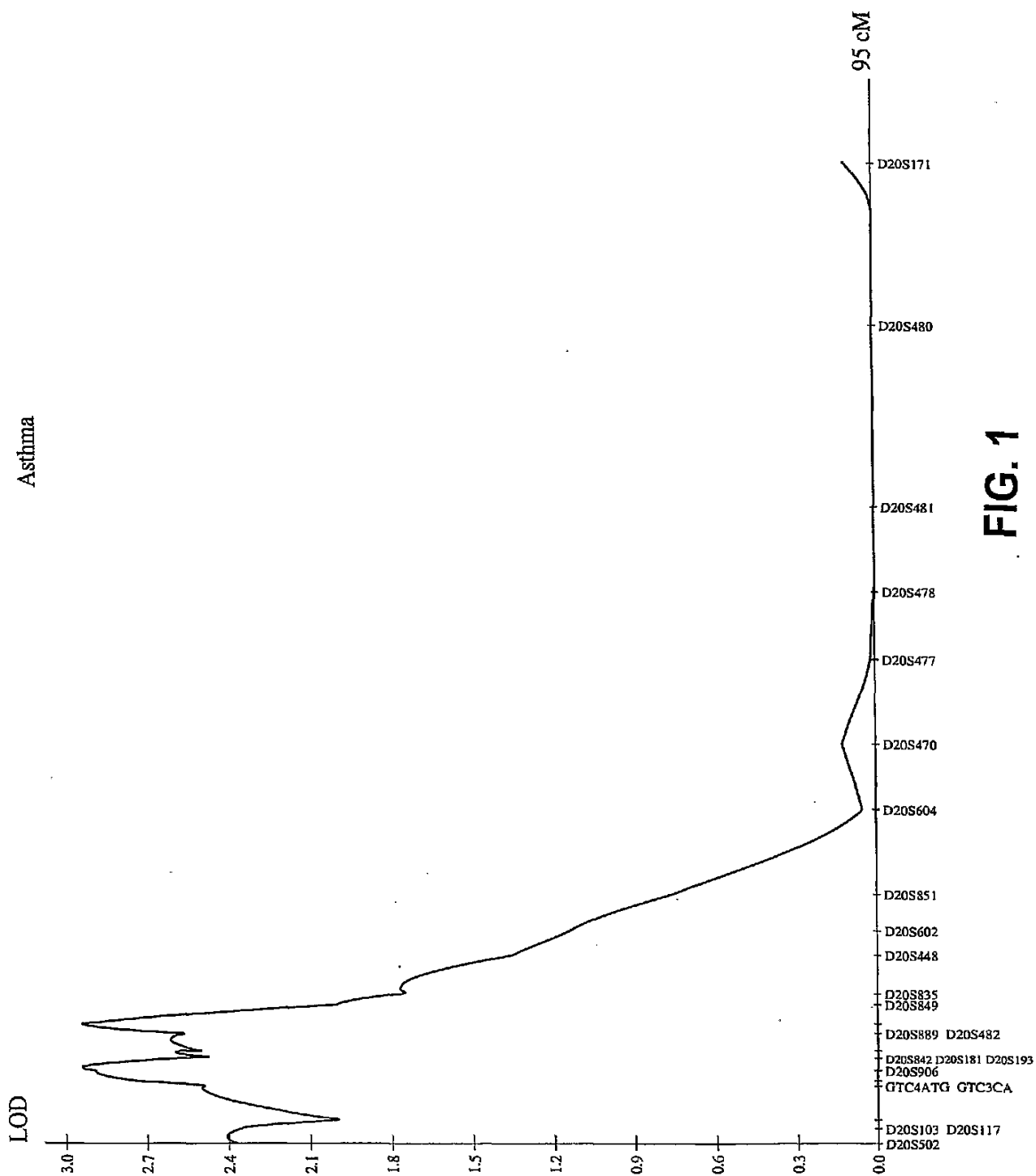
b. detecting the hybridization complex of (a), wherein detection of the complex determines the pharmacogenetic profile.

83. A method of determining a pharmacogenetic profile comprising:

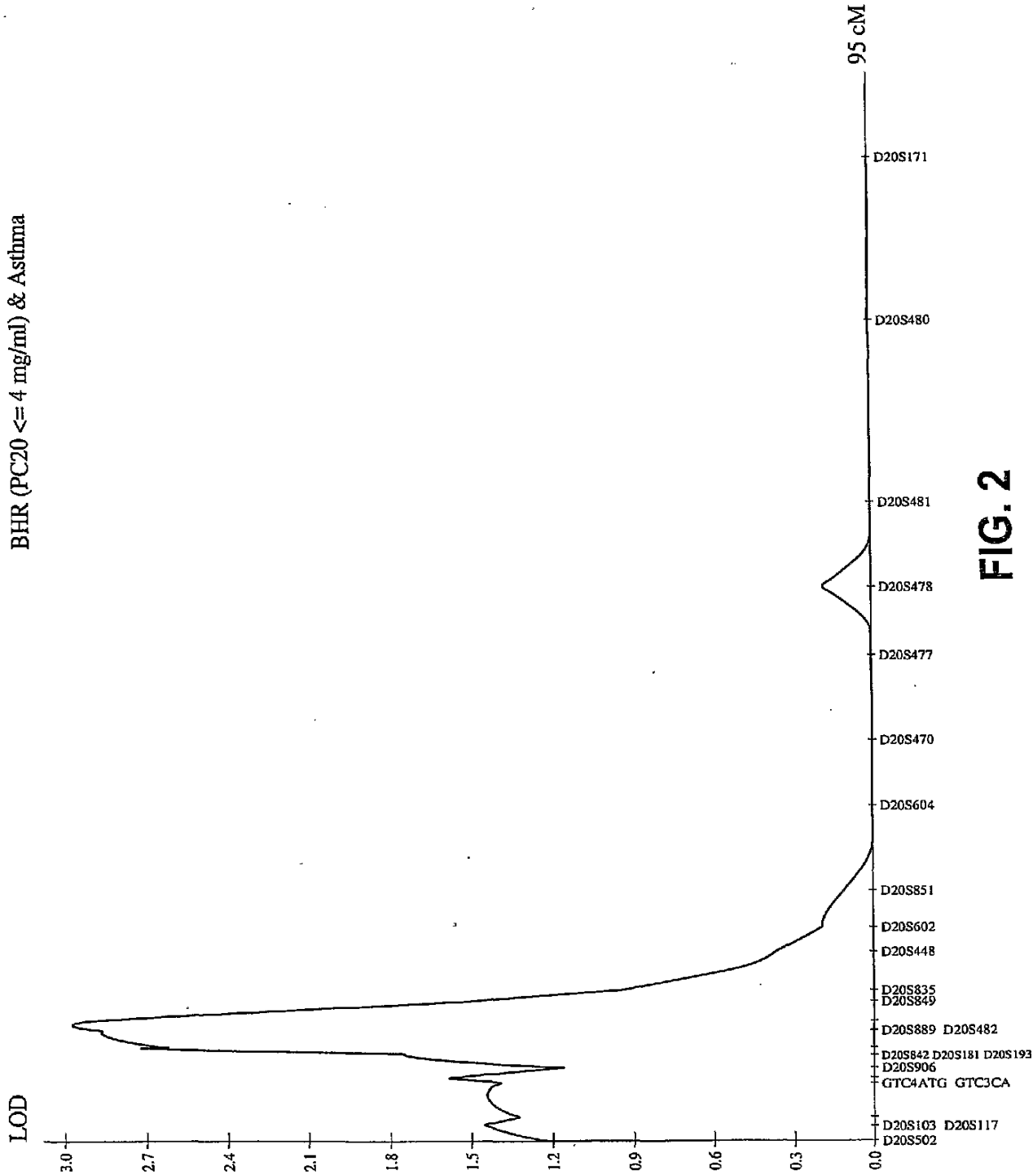
- 5       a. contacting the isolated antibody of claim 41 with a biological sample under high stringency conditions that allow the antibody to hybridize to an amino acid sequence in the sample, and thereby form a complex; and
- b. detecting the complex of (a), wherein detection of the complex determines the pharmacogenetic profile.

10       84. A cell line comprising the isolated nucleic acid of any one of claims 8, 19, 26, and 28.

      85. A biochip comprising the isolated nucleic acid of any one of claims 8, 19, 26, and 28.



BHR (PC20  $\leq$  4 mg/ml) & Asthma



BHR (PC20  $\leq$  16 mg/ml) & Asthma

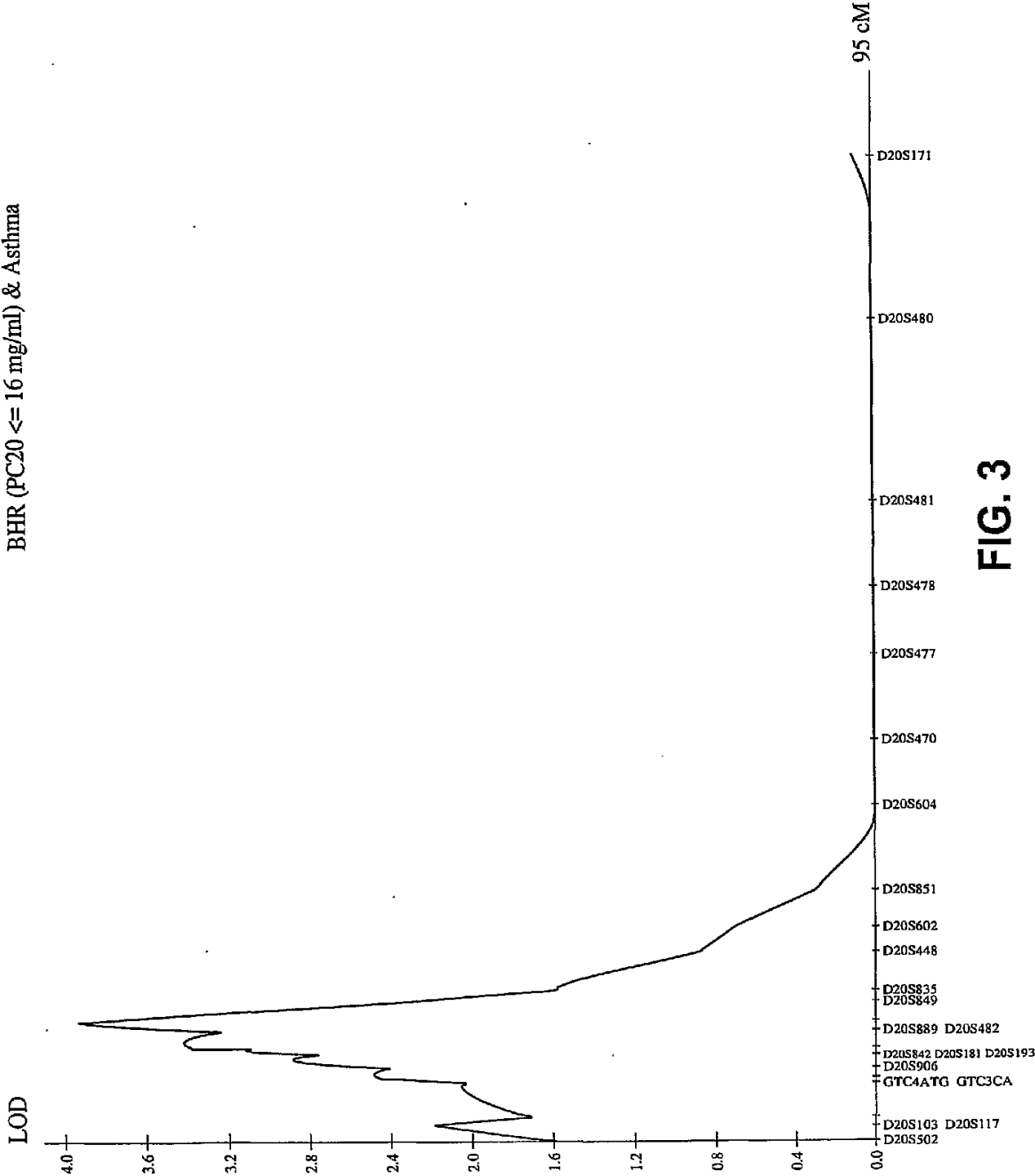


FIG. 3



High Total IgE & Asthma

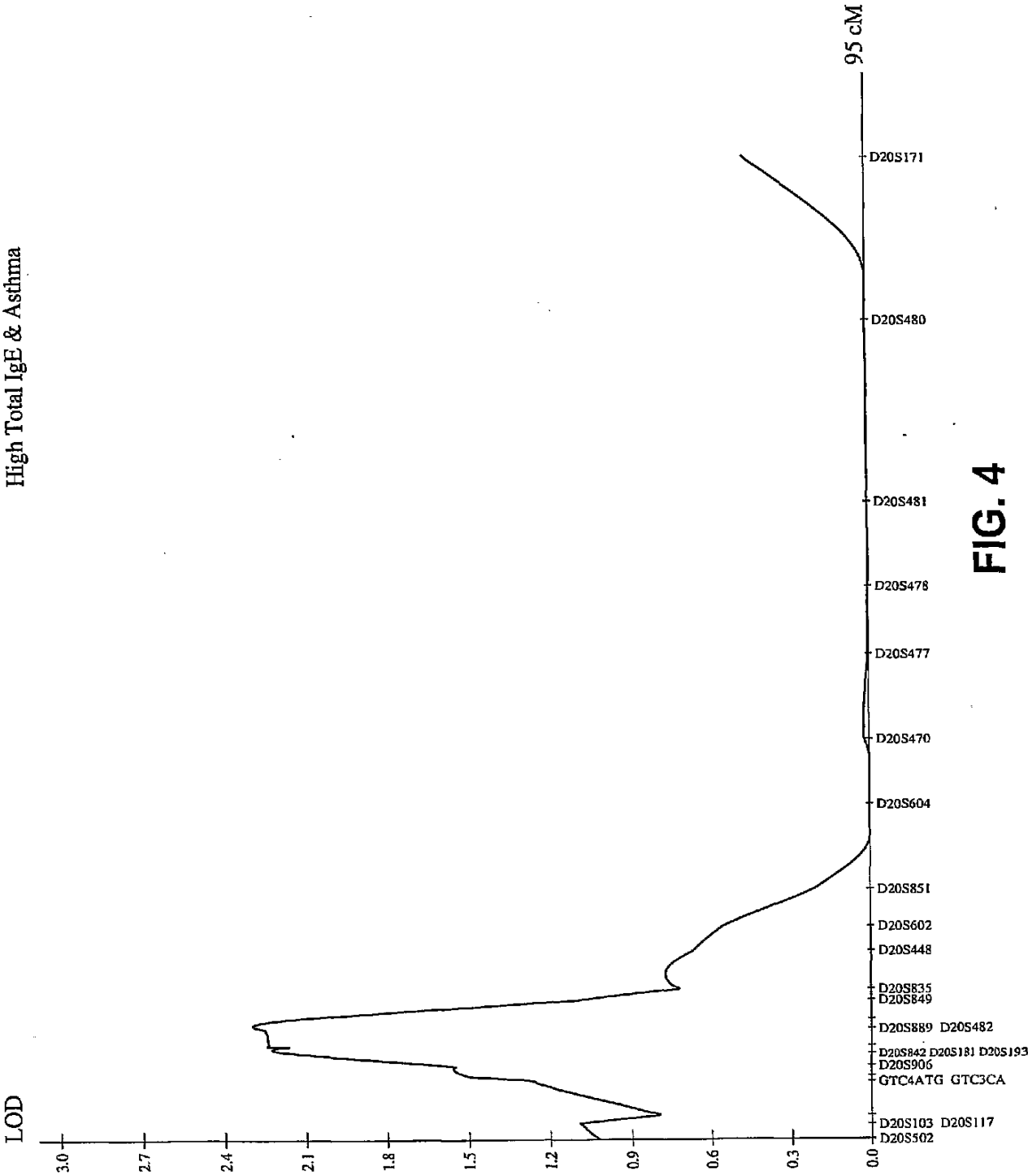


FIG. 4

High Specific IgE & Asthma

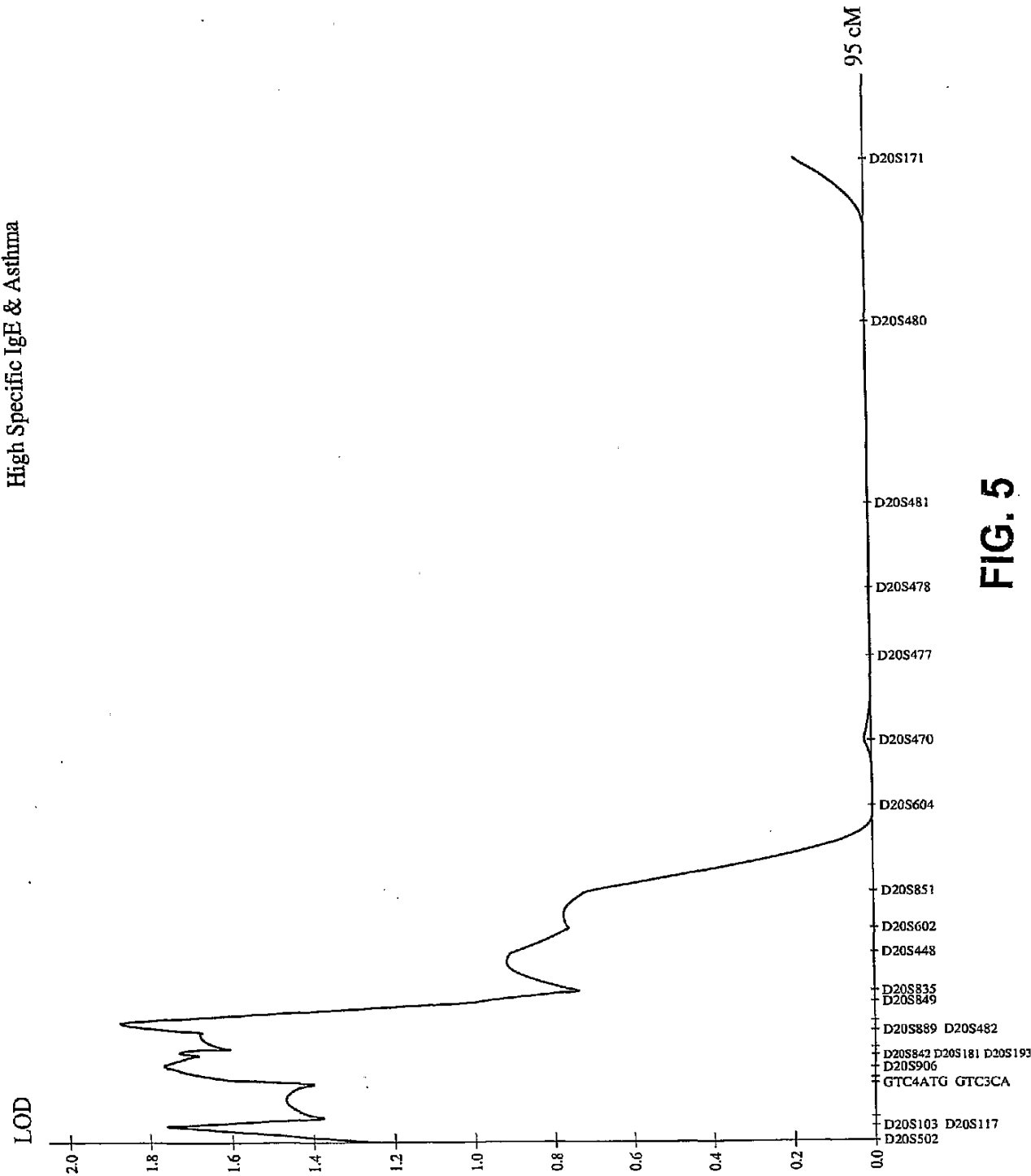


FIG. 5

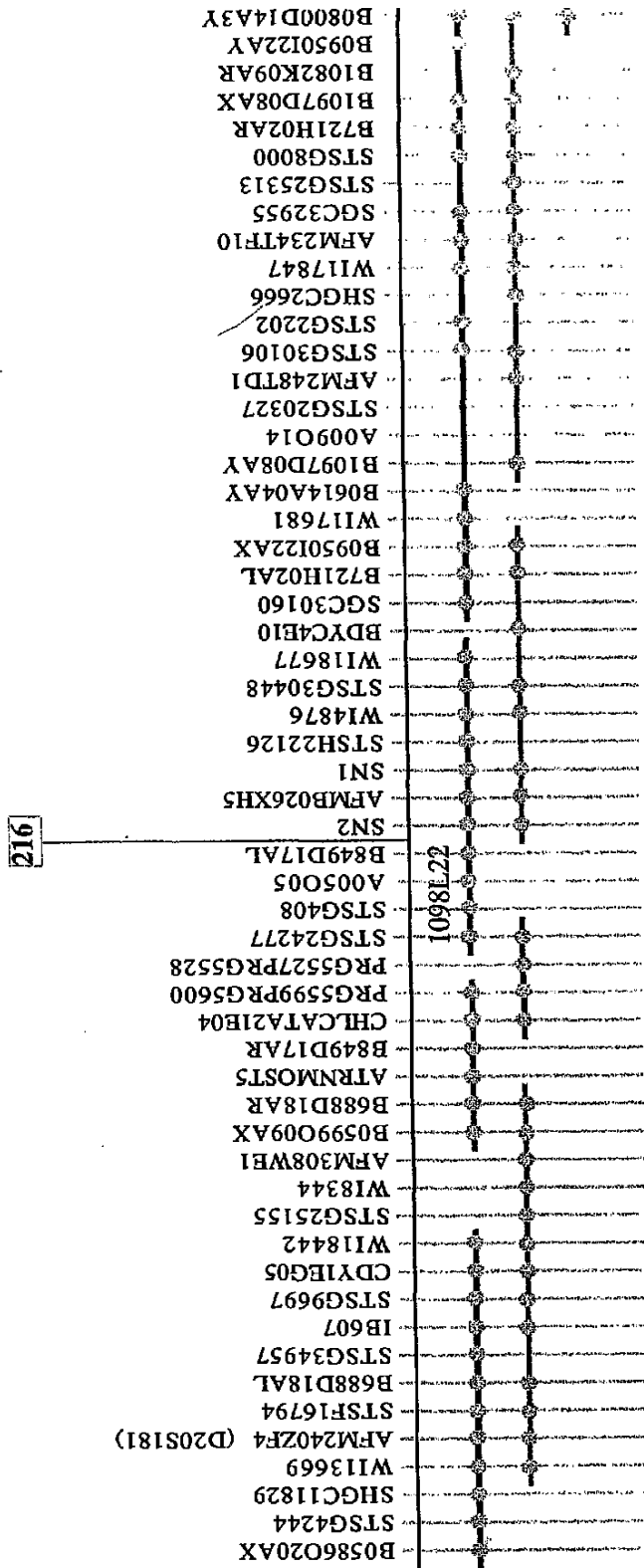


FIG. 6

B0624E22A2X  
B0870L06A2Y  
W13772  
B768F06AL  
W12640  
STSG25189  
RM8ARM8B  
APM036YA3  
UT235  
STSG25337  
UT11704  
CHLCGATA149E11  
CHLCGATA51D03 (D20S482)  
B819A20AL  
B583H11AR  
B715K24AR  
B0934N19A2X  
B1122K02A2Y  
B736H10AR  
STSG20232  
STSG20381  
W118738  
PDG3PDG6  
EST498946  
W17784  
W13651  
SGC34960  
B727P10AR  
B583H11AL  
B819A20AR  
B715K24AL  
B772K01AR  
B1122K02A2X  
B780N09AL  
B736H10AL  
B602E18AL  
B1033I16AY  
STSR73406  
B0686P06AY  
B732A06AR  
SHGC16916  
R79078  
SGC35090  
W111626  
STSG8643  
STSG5017  
AFMB352XD9 (D20S895)  
STSG42745  
STSG10203  
B602E18AR

FIG. 6

>BAC1098L22 sequence

**FIG. 7**

9/157

gggggagtaattgtaacatgtatatgtgtacataaatctggtgacatcagtagattatttccataacctgttttaacctctgga  
tctctgctaggggagaaagagaggtcactgataattagctaggttggtattagccacctgagttccttggagtttaaggtat  
tataatagtgcataagactgtataattaccactaagaagtgtagatctcagctggatgtggtggctcacacctgtaatcc  
cagcactttgggaggctgaggtgggtggatcgctgaggtcaggagtccaagaccagcctgaccaatatagtaaaacccc  
gtctctactaaaaatacaaaaatttagccaggctaggtgggtgtgctgttaactccagctactcaggaggccgagggcagg  
agaattgcttgaaccagcagggtggaggtggcagtgagctgagattgcgccacgcactccagctgggtgacagagcga  
tactcgtctcaaaacaaaaaagaacagcaaaaaaaagaattgtacatctccttgggtcttccatagcccagccatctc  
aaaaagaagagagcaccttcttgtcaagagtttctaagcccagaaagggtcaagttctctgcttgtccacccagtgctctc  
agggggcttatagtcaatattccatgatcacatttttgtcatttttagtctggagtcataaattgtgatcctaccgtcagt  
taagttagactcatagaacaaagctcttccagcagtttcagctgtggtacagaaacggttagtggaattgttctaccaag  
caggggaggagttgagggcaacacttcttgggtacagcctcctcatgtgaaggtatggaaattggcctgggtctcggc  
tgctcttgggcctctgtgtaccacctatagggccattctgagactggtaggaggtgcctctgtatttagtttctccaatt  
gtccctttttcagtgcaacttagatgggttagtgacactcaaacattgggtgacataattcttagtgtgtttacctcaggc  
tactgtgaccacatttggtatttcataatattttgatagcttttccagatttcagaatctccattggtgactgtctctgt  
tgtttctctttccatgtccaaatgtgggtctcttccagcattccatccttggctggcagctgacctttcccatagttatt  
cactcctcagaaaatggatggcaccagcctttgcttattaccctgggtgtcttaacaattgactctcgagtccagcgag  
ttaaaagggttcaaccagggtggccacagatatacttctgtaccctttctctctcttaggttttctaactctaaacgt  
ttctgggtattctaattctgctgtggccacggtttatgaacagaaaattccagctcttaattgataagactgacagatgagag  
acaactgaagtgaatgtccttccacagctatacctctagatgtagcccagtttagtagagcccagatttttatggaaaaa  
caagaaaggacacctagcctaacccttaggacagagggtgggtggtagaagctggaaggaggtgacccctgcaggtgag  
aaggagtgaattaggtatgtcgaagacggaagggtttctgtgatttaattagtgtcccccatctgtgagatgtagagggag  
atgattaaaggagtggtctctttagtgagctgcaggttaagtttgcatggttggtgcaggctggatgggaggggatttt  
atagttgagcctcaggaaggaacacaggccagacactgcgaggcactggtacaattagtaaatggatttgtaattgtatcgt  
gaaggcaagaaatattgaaggcctttaaaggatatttttaatttgtctttctttcttttggtatctgtttatttttagg  
tcttttagtaggcatgtgtatttttctctcaaaaatggaaataggctgggtgtggtggctcctgctgtaatcccagc  
attttgggaggccagtgagggaggatttgcttgagcccaggagttcaagactagcctgggcaacacaggggagacctcgtct  
ctacaaagggaagttttttaagaatttagccgggcatgattggtggcacatacttgtagttccggctacttgggaggctga  
ggcaagagggttgcttgagcccagggtttgaggtctgagttagccatgatcatgctcctgacactcagcatgggcaaca  
gagcgagacctgtctcaaaagaaagcaaggagggaatgacagtaattcttttgttttaactacaaaaattaggaa  
tacttaccatttcttggctaaactttatatatttgatttttaaaacttgttaaaaattgcaatgagaaggaaatttcagga  
gagcagaagacagactgtcccagggtgtcactgtcctattattccctataaaaatccagtgccaggatggatgaatggataa  
agcaaatgtggtataagttgaatatcccttatcaaaaatgctttggacgaggagtggttcgaattttagaatatatttgcat  
tatactaaccagttaagcatccctaattctggaaatccaaaatgctctagtgaacatttctcctcagcatccatcatgtg  
gcactcaaaaagtttttagattttggagtggtttggatttcagataggagactcagcctgtgtttgggggtagccatct  
ctcatatagacatttcagaacttaaatattgtcttgcataatttctgtgaatttttgatatattatctctctgagct  
acatttttatctttataaaaatggccatattgaagtgatgatctatcctaattccatggctgagtcagggtataaaga  
gggtttctctgtgtctgtgggtatacttaacttggtggtttttatctagaagcttggttttggtcaagatgttggttatat  
tcaggccaggcatggtggctcatgctataatctcaacattttgggaggccaagggtgggaggatcacttgagctcaggag  
tttgaactagccaggggcaacatggcaagactccatctccaaaatttaattaaaaaagatactactctgtattcatagt  
tgtgtctcttttgcttttagtccaagctcaecttaacccctgggtggcctcggaagatcaatgtgtcctactggtgc  
tgggaagatatgtccccatttacaactagtttactacagtggtatgcgtctgagcccagtgatgctggattctgtggaat  
tttatcagaacccagtaactcggggactgaaggctgcaacctgcatcaacccactcaatggtagtgctgtgtgaaaggcctg  
gtaagttcacagggtgaattaggtggtattcagagttttatgtgagagaaacataggaggcatagttcattgctgagatg  
tgtgaagttagtcatgaaaacagatgaagtattgatttcaagcatgcaaaagaagatataaccagatttcagaagcagaa  
ggaaatattctgggacctgaatagtttaattataagcaaaactaaaaaactaacactactcgaaagaactgatatt  
ctaattaaacatgagattgtaggtttatttgaccagaaaaagtattgagaattgctctgaaaagcaaaattattgggtgtt  
ggcagagaaatgctgtggaagaagaacaaaaaggaaaaataaaacaaagaaaagatttagtaagcaagcaagtgactg  
caggggacagtggtcagaaaggtagtgtcaacaggggagaaaaatgatgagagcagttctgtgaagcaggacagaagcaacc  
cagaactacgcgagagctgcaggagagtagtgagcagacagacactcggagtagcttcccagctttcagctctctctgggc  
tacatttttggtctaactaaaggcaggccaggggtagctgctgcagcatgaagacaactcaaaataagcctcctcctctg  
cagagggactcacaagcagtccgttggaaattgtgggtattttgggaggtggactccatttgcatgtgtccaaattttgg  
gaagtaattttgtctgaggaattacagaggtatagagagaaattcatagcatgtagatttctaataattgattttctaa  
cttctctaggggagctgccccagcctttcagaattctgaatccttttgttgaacttaatgaatgccgtagaccttcttc

FIG. 7

10/157

cctccaaaattgcaaacatgagattttacatattcaagtggtattgtaaaacccctaaagttcatcctgggtcccaaat  
ctaaggggcttacagccctacttttatggaaagattcactgtttcctcatctttgtccttaatgctctttgaaaagaaaa  
tttattttttccacaagtgtgtaaatacacttgactaaacagtgacttgatgacttttatgttttctattttctcatt  
taattgatttcatggcacattaatgggtcatcaccacatttgaaaagctctggctgggcacagtggtcaogcctgtaa  
tcccagcactttaggaggccaagacgggtggatcacaaggtcaagatatcaagaccatcctggctaacatagtgaacccc  
catctctactaaaaatacaaaaaattagccagacgtgatggcgacccctgtagtcccagctacttgggaggctgaggca  
ggagaatcacttgaacccggggaggcggaggttgagtgagccgaaattgtgcccactgcactctagcctggagacagagt  
agactccgtctcaaaaaaaagaaaagaaaagaaaagtgctttgacctagcggacatggtgggagctctaagt  
tctctctgggtttcattcccagcaaacacagtgctaaagcagtgccggacaccatgtgccttgaggacagcatgtggag  
attgaccagcggcagctctgagtgcatgtggtgcagcaacatgaagcagtggtggaactccaatgcctacgtggcctcc  
ttcccttttggccagtgatggaatggtatagatgagcacctgccccgtgaagtgaaaaaggagccctaggcacttat  
gcatgccccctgtataggcaaccaactcagccatgaggtgtgctgcagcctctgaacattttagaacaagactggaca  
tgacctctgctcaaacctgaccagagactgccatcgagacctgtgctgcctattgagaaccttcatacagaatcaggcaca  
ttgacagtaaaataatgtgaagatagatcacagagtacagaaaataacttgtccaacttcagtgctcatattgtcctaaccat  
gtaatatctccatctgaatttcttaatttgtaaagttaattgctttccaatagataatctctaaggtcccttttgccct  
caacatcctgggttgagagaggagggaaggggtcatctctgttatgtattgggcaaaaatactgggctctttacattcatt  
atctcttttaataatcaagacagaataatatttttgactcaagccagtgaatagtctgttataaaaaaaagttaataca  
gtgaattcagatctacctgtgatagtcattgcaacttttttttttaataagctgaaaattgttcaggctactgtacctg  
tagtcattgcttgagcaaccaggtgtggtgtgactgatcccagcaatactggcaaagggaaatgcatagagggtt  
cctataaaggaccagtgaagatgccttcgcaagccctacaggaaatttctatccacagccccctgctcaattccagcatg  
tgtctagaggacagcagatataactgggtctttcattcactgtccaggtgaagatgccttgcatatccaaattcaagtgttt  
cactactgatttatgaagaataaaaaccttgaaagctacgttgtgtatgttaactccctgcctcagcccccttccctcc  
tcctaattggttggtacaagaaggaatagaccagaagctgggtccaggccctgacctggacctgctgagagtgggtgggt  
tcctaagaaaaccaattctaagaaattggcctttgattcagactgaagtgaacctcagcaatgtgtctgtgggtttcta  
gaacagttgggagaggctgggtgggtgcaagactcctcagagattagcagtcagaacttctctaagagcctgccattg  
acaacagggctgtttgtgaggactttgtgaagggaaagtcactgtaaacaaagctaaaagggcagagacagactgggaga  
aaatacctgcactgcatgtaacaactgatgatcatccagaatatgtaaactcccacctccagggaaaaaataagaagct  
aatgtgaacccaacagaaaaagtggttatgtgagataaagaagccattctcagaaaaggaaatagaacaggaaatgta  
aactaacaccagggaaccaattttttgtctactaaactggatcaaattttccgtgttccctttttccataactaagatatg  
ggggaacctgcaattctatttttaattctgctaggagttaactttttaacgaaatatttaaatctctgctttttcatgaata  
tcacgaatatatctggtaaaatgacaaccagagaatgggagaaaatatttgcaaagtatatatctaataagaatccaat  
gtccagaacacgttaactcttaaaactcaacaatagaagacaaccaatgaagaatggataaaggatttaaatagacat  
atgtccagagaagaaatacaaatggccaataagcacatgaaaagatactcaatatcattcatcattaccaagaatgcaa  
gtcaaagccacaatgagataccactttacaccactgagatggctgtaatacaataaaacaggtaataacaggtattggaa  
ggatatgtagaaattggaactctcatgcaggttggcagggtccaaaaagttaaatatagagttatcatacggcccagca  
gttttactcctagggtacatacccaagaaaattgaaaacatattgttacccaaaaacttgtatataaatgcttatgcttata  
gcatcattctctatagcatcatgctcaaagcaacattattcataataagcaaaaagtgaaaacaagccaaatacctgttag  
ctggcaaatggatgaacaaagtgttgatatctcatacagtgtaatgttatttggcaataaaaagggaatgaagtactgaca  
tattttctgacatggattaccctaaaaaacatgctatgtaaaagaagccaggtgcaaaagattatgcgttgcatgatgc  
catttatatgaaatgtccaaagacagaaagtagatgtttagtgggttccctagggtgggcatggttaatgaagaataatag  
ggatagggttttctagagttagctgcagcatttttctgacatggattaccctaaaaaaatgctatgtgaaagaagcca  
ggtaaaaaagattatgcattgcatgatgccatttatatgaaatgtccaaaggcagaaagtagatgtttagtgggttcccta  
gggctggggatggaaacaaggaataataggcatgaggttttctagagttagctgcagttatttttctgacatgaattacc  
ctaaaaaacatgctatgtaaaagaagccaggtacaaaagattatgcattgcatgatgccatttatatgaaatgtccaaa  
ggcagaaagtagatgtttagtgggttccctagggtggggatggaaacaaggaataataggcatgagattttgtagagtag  
ctgcaccttttatgctctcaccaacatcatatgaaggttcatactcttccacatccttgccaatacttgttactatctt  
tttaattgaaaactgttctagtggatgagaaatggcactctcactgttgttttctatttgtattttcctgataactaatgaaa  
ttgaacatcaacttcattagttagccttttgatatcttctttggagaaatattttagtcaaatcttttgccatttttca  
gttatgttgtcttttttattattgagttgtaagagttcttttatgaattctgaggtcccttatggatatatttttgcaa  
atactttctccatttatatgggtttcttttcaactttcttaatatgccttttgaaatcaaaagttttcagttttgattaa  
ttccatgtatcagtggtttttttatcccatgtgctttttgggtattgtatctagaaaaatcagtgcccaagtaaccaggga  
tcacatagatttgctcctaagttttctttgaaaagtttcataagatttggtgttacattaggtcccttagtcattttgag  
ttaattttttgtgtatggtgtgaggttaggggtccaaactctcttttgcctgtggatgtgggtcctgcaccattttgtgaaa

FIG. 7

11/157

agatttttttttttttttaaccattgaattttcatggcacatttgggtgcagatcagttgactgtgaatctaggaactttt  
ttctaagctctcatttttgtgcagttgaccaatgttttctccttttgccaattattacactgttttgattactgtggctttg  
tataagttttaaatttgggaaggctaaagtctccaccttggttcttattttgcaagactgttacagctattctgggtttc  
ttgcatgttcatatttaattttaggatgagcttgccattttctgcaacaacaaaaagccaactgggttgataaggtttgca  
ttgaatctactgaccaatttgggggtgtattgctgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgagagagggto  
tcaactctgtcaccaggctggagcgcaggtgggtgtatcttggtcactctgcagctccgctttcccgagctcaagtgattc  
tccagctcagcctctcgagttaactggtaactacaggcatgagctaccaacaccagctaatttttatatttttttagaga  
cagggtttcgccatggtgccaggtgggtcttgaactcctgagctcaaagcaatcctcctggcctttgcctccctagttc  
tgggattacaggtgtgagccaccacgcctgggccccttgccattttaacaagggttaatctctagcctgtgaacctgga  
atgtcaatttattaccctcttactttcttttggcattgttttatagttttcagtgatcatctgtctgttcttctttt  
ttaggggaaaaataattcagtccttataaataatgatattgttttagtttttttatacatgccctttatgggttgagaa  
agtttctctctattctagtttttgggggttttgaatgaatgagtgctagatttttgcacacacttttctgtgtctac  
ttagatgatcacgaagttttgtctcttataatactatgtattacattaattaagaatgttaaacacagttttgcatctct  
gggataaaagccatttagtcatgatgtatacttttttatcatgatgtatatttttaactctgctgttggtattctattgct  
agttggggattttgtgtgtacattcatgagggggatataagtttgtgggttttcttcttttgtgatctactatcgtgtttc  
atagcagggtagtactggcctcctagagtaagttggaacgtattcctcctcctctatttactggaagagttcgtcaaggga  
ttagcattaattctctctttagatgttgattgaattcaccagtgaaacaaatgggcccctggccttttctttacaggaaaa  
tttttaattattatctcaatctgtttgttatagatctattcagattttcttttctgagtgagtcagttttggtaatctg  
ttttcttaggaatttgtctattctcagtaaatcaaatttgttaacatacagttcgtagatttttcttttggttttttttt  
tttttttggagacagagtcctcactcttccaccagtcctggagtgcagtgacgtgatctctgcttactgcaacctccgct  
ccagggttcaagtgattctcctgctcagcccccaagtagctgggactacagggcgacgccaccactcctagctaat  
tttattttgttttgttttggctttttttagtagagacaggggtttaccatgtagaccaggttggtctcgatctcctgacct  
cgtgatctgctgcttggcctcccaagtgctgggattacaggcgtgagccaacgcgccggcgagttcatagttattt  
cttgagggtgtctctttagcaaatatgtagtctcctgttttaacctgttaacgattttttttctgtctaatttgggtttat  
tattcttttacattttgcaagggtctaaatatgtgaggggttttttttttaagagcttgctctcttatattgtggatg  
caataatttaagccttattgaatgtactaaaattcgtatttttctctctcatatttcttgcattaacctctccttcag  
gatttgttgctgtgcgtgtttatcctgggtggatttcggagattgagagttttcttatactgggttaactcctcaaggtta  
ttgtattataaactaaatgtgtagatctgtcaatatttggtcagtggttgggtgtgctgagctgttgtgaaagtggtgat  
gatgtcctggaaaagctagcaagcagcctcaggacactcagctcgttgcacttcgactgttgtgactgggtgtgcgc  
atctcctcattctggaaaggaaggagcaatggacctgggtgggtgggtgcaattagcatgttcatgggtgtggcaagt  
cctccagccttgggtggtaggggccacctaaagatagctggcacattgccttgagttttcttctcattctgtctgttgc  
ttgcttttatttttgagatggagtttcaactctgttgccagggtggagtgcattgggtttcttggctcactgcaacc  
tctgctcccggttcaagcggttctcctgctcagcctcccaagtgagctgggattacagggcactgccaccacactctgg  
ctaatttttgtatttttagtagagacaggggttcaactgttgccaggctgatctcgaactcctgacctcaggtgatc  
caactccttggcctcccaagtgatgggattacaggcatgagccaccacaccagctgagttttcttcaaacactcaa  
ataactaacagggtgcataaacagaacaagtacaaggctatggaagggtgccaagatggtgaaaaactgcatcacactgatc  
aaccaaggcagatctgaagaatcagattttgcttacagtgcagggtgtgcactatgtaaatgtttccccatcctatcc  
cccctagagttctgcagtgacagcctgtgcagctgtcctgagcagctcctcagaggtccaagctcatatcctttctag  
ttgtaaagccctatatccattaggagtccttagtaatcccactgtggaaatgttttctgtaaaagatgggtttttctta  
ataatgaaaaatgtaaagcaaatataattagaatatgaaggaatgggttaataactgttaagattcatcaaatcatt  
ctaccatttgaagttttgtttataattgatgaaataagggtgtaaaatagttgtgatctaagtttaagaaagggtgaatatt  
aatatgaacataattgtgcaaattatgttctctgggtatagaaaaataccagaagaaaaatatcaaatgtgaacagtggt  
tctctgagtggttggttggtatgatcttcccatcctctctattttttaagtatactgtcatgagcatacgtttttt  
ttagtgataaacaacaaaaaatatttttaatatcactatgggctttcacttgctgtatcattttcaattttctctgg  
actagctcttctttgatatttttctattttctcaattagttttttagtggccagttttgtgtatttatcaagata  
aattaactacagagatagttgagataaatttattgaacttaactacatgcagttcttccatgggtacagacccttgaa  
actcttttctgtcagcttgccaatgcaacgggccacagtaaatgcatcaatcagagcatctgtgagaagtggtgagaacct  
gaccacagggaagcactgcgagacctgcatatctggcttctacggtgatccaccaatggagggaaatgtcagcgttaagt  
caaattgggtcaggttactcatggcaaatcggtgtggaacacagcactctatttgacttgaatatctaggaggaaaaaa  
gccactttgtttgggataccacatttcttataaatcatactggtcagaagctcagctgagctggaagagaagcaacagg  
tttcagtgtagctgggttaagcaacaacgcaaaaatgttagagccaagcttgacacactctgcagtggtataagcaa  
gggcccacaaacagagaagaaagttgccaacttccctgtttctcttcccttgaacaggggcacaggtcaatagtcagagg  
gatcagctcagacatgggagtccttggggctcactcactgaggccggacaaaagggtccagcagtatatggacactaag

FIG. 7



12/157

gttggtgggaagggcaaggtgtaagggctaggactgggaaagtctctgggttgaggtcagaggttaggaaagtgtcttctgaa  
ttccccctcctcccccatcaggaggccttgggtgcctgaagaagacctcagcccaaggcctcagatagggagcctgggaa  
taaaggctccagggttaggaatgtgaacatgtgaaggagcgtgacccgtgaggttagacctgagtccttaactgcacctct  
ctcagagccccacacttagggccacctcacaggtcatgctctgccaggtaaagcctgtggcatgacctgtgggtcaggccc  
tgaaaaatcacacacgggctttttaaccaggagccagactcctcaatacttatagtgatttttaaagggttaaatagtc  
cgtgttggttggtgttagtggtcacgcctgtaatcccagcactttgggagggccaaggcaggtggatcacttgagggtta  
aggagtttgagaccagcctgaccaacatggtgaaaccccatctctactaaaaatagaaaaattagctgggcatggtggcg  
ggggcctgtaatccaagccactcaggaggctgaggcatgagaattgcttgaacctgggagggcgagggttgacgtgagcta  
agatcacgccactgcaactccagcctgggcaacagagtgagactcagtcctcaaaaaaaaaaaaaagaaaaaaaaaaaaagt  
cctgtgtttcccatttatgtcaaatagaagcctgcaagcaataggacattttattatgagaaacaaaaattatgacaagt  
gatattaatgaactgcttttttattctagttccagccaaaaacataaatgagtaactatagtcgaaagacttttcaaag  
ttctgagcaggaaaagtaaacacatagggaaatctctacactcccagatcccaactccaaactcttggcatggcc  
cttcagaatctgatctgatccaagaccagccaggagagaggcagcgttctagctgtgtgtcccttactctccctggggtt  
cagtttctcctaataatagggtgggtgttaagatgaaagcagagcagcacatgacacatggtccacaggatactagtc  
atacttagtagatgtgggcaattcttattacccctctaatgccattctccactactcctaagtaaaagtctctgtctcaga  
ttaaagggtatttgggtgaaaatgttttcccatcttctgtgaatagacttaatccgttagacagtaaccagagtaacctaa  
cagagtggtggagtgtaacaggcattgtgtctgttaccactcaagacacaggtgcttttattatccccaaagatcaagtaa  
actgcccagggtctcacagtgagatgtggccaagttaagaacacagccagctgtgtgtgaccttaattacttgggtggt  
tgtgtctccagatgccatcataccacatcagcagctgctatctagaagcatcacattttctctgtgagatctacaggg  
cctgagtacaatttgcttatttttccagagtcctaataccaggtagaaagcgtgggtgtttatagaatactgcagatggca  
ccaagtctttgatgttttcttcttaaaattgtatagtccttaggtaacgatagtagccatgatcttctgaatgcttactg  
tgtttgagacatcatatacatatcatcacaaacacccctatgagataggtaactgttgttatcccatcatatggacaagg  
aagctgaaactcagaaagggttaagtagcatccccagagtgacagaagagccaaagtagaggacgtgaattcaaaacccagg  
cagatggacgtaaaggcctttgcttttcaatatacatattacacaccttccccagctctcaacagagccaagtgtcagacact  
catgattctgacttcagcatccattctgttaatctctgcactattctgtttttatagttttcttgggtgatccaaagga  
tttatttttaagcatacctctctccagctgaatgcctttcatttatccaccagcaaacagggtataaagtgaagggt  
gttgacaaaaaggcctttcacttttttcaactggagtataatatttatcacggcttgtattacgttggatgataaaagga  
gagatgttgagttggccagatgagggagatgggaagagctatttctgttaaaggcaaggaaaagggtgaagcaagggtgg  
catgtcctgagggcccttaagcgctgaccttggcattggatcctgtgtctgagccctgcggtgattccaaatgtgacca  
tcgtggagaaggcacctcagccaagaccagccctcctcagaggacagtcacctccagagcacctccctaggggcagcag  
ctatgtctgtgtccccacaccagcacctggcattctagagatgcttcacaaatatttatgaatgaatcaagaatcat  
agcagtgaaaaagagagctcattgaaaagatcaaaatagtcattgcttcagaaggcagtggggtaccagcattcactttcc  
tctgtcctttgtttgggttcccttaattgttacctatacattattatcattagctgaaaaatcatgaacttaccattga  
atgctgtacttttaactgtcttctcactgaactcaattcagataaaaattgcttgggtttgggaaaagtctccaatagtc  
agaattttttaagccagttgtgacctctgttctcttttctctctgcagcatgcaagtgcattgggcacgcgtctctgtg  
caacaccaacacgggcaagtgtctctgcaccaccaaggggcgtcaagggggaagagtgccagctgtgagtaccatactccc  
tggaaccaccaggggaggaccaagaggctgtgcagctgacctgaacccacccctgagagccacccacttccctgtgtctgtt  
gctgtgggctctgagggatccctgggttgatttagtttgaaattttgcccattctatttcagacagggtcagtccccaaat  
gaggaggtcgtcaggttagcagcaattccttaattggctcttgaaattcacatttttggtttaaatgatactgacatttccctg  
ggtgtccatttgagtagtcatttttaacttcagcaactacttgatttttgctatgtcaagagattatactcttcccaa  
agagtagagatgggaagagctgggtgtttgggtgatggcatttatttggcgtttgggtgtctgatgtggaaatgatcctct  
tacctttaacatttcccatctactctagcattttctgttgaagcatgtcaggcatcttctggagaggggtttgaagtcc  
ccatccatgaggatgcaggtgcagcagcatcaggcatgtttaagggtttcagacctggccctgctccactgagcaagggtg  
ctcttgaggcaaatcactgttcccccatctgtcagtagtgaggactagcagtaattgttttgggtgtcctggtaggttga  
aacaggaaggatagttcctcaatcaatgtataccgggctgtctcactgggaaacctcataaaatgcatgtatccgcattt  
actcatttctcaagaaaaaaacttaaaaagtgttttagtgtcaaaacactgagctgattaatcattgttaacatcattatttc  
ttaaataacagtagtaataatgccaccagtgattgagacatcattgtctgggtgctactcagcaggggtttgggtctgtctact  
tttagtttatatctgcataatgtttactagcttcatgttgccctaaagccaatgtgaagaaaaagtctaataccaagg  
ttaaatttaacttttaggtccacattcccttaaacagaaaaatgtcatcgagcaaaatataactccttaccctcagttta  
gctgatgcaacaaagacatgaggtggcatcaagtaagactttcaagttccttcgagatcgagatgtaggcaggtgctgc  
tcatccctggctaaagggacacagcatacctgctgggaaagggcattaccacttccgccttctcctctcatgca  
tagctgttggtttacctaattgcttcatgaagtgagggtgggttttgctgtgtttatagatgatctgtatggggaa  
attgttttctgagtagaacatatttttttctacagactgggtgaacactgttccagggtttaaaaaacagtgattctt

FIG. 7

13/157

ggatatttagtcttctcacttgtgtttaagaagagaattagtttatccagggaaacataggaaagaagtggagaaattag  
tatttgaaaaatgttttgactgcacttttagaagaatatgtagtccacacaaaattgaaaatgttttctattatagtaaa  
taggtttattacatttggatttctctcattgctttggcaaatgcttgtcaattgtcctcatatagttgggtgtgtgaactg  
ctgagcagtcagttatgaagcgttcatgcatgctcttcttagactgctgagctgagttatgggtgaaggatgcaattata  
atggctccagactatctgtactttatataaaaggatctgtttgggttttaaaatgaattcagtttctgttttaaaatagcag  
tataaaatagtccttttcccccagatgtgaggtagaaaatogataccaaggaaaccctctcagaggaaacatgttatt  
gtaagtgggttttgcgaattcttatttctagaagcaaagtagttcagtaaaacttcattgtttaaacgggtttgagaatagt  
aagtgtatataagactatagcagccaccaatgaagtgttccagacttgatatgtttacattctgttaagtttactacata  
taggagcactgttttaagctgttttaattgtgtttgggttaacgttaatgtgtccatagcagatagcaggagagagtag  
agaggcatgcatctttgtctatccacatttatgttctcttaaaactttactttatttgccattacctagtgggggtcatc  
atatttctgtgttttaggatgtagatcaaaaacagaattcttacaatatgggtgaacttttctcattctgtctgggaactg  
atltgtgttaaccaccttgaggtgaaaggtgagctgacaaagtgaccatgtttttatgggtaaatgtgttttctcttta  
tgggagaatccacatggttagaccagagtacgggaccagaaaaaaggagttaatgttatggcatatccattgtaatatat  
acctgctgtactgggttgaataacatcccttcaaaattcatgtccacttggaacctgtgaatatgacttcatatatatat  
atataaaacttctaaaagaaaaaagagaaaaaacatcttgaccttgggtgggaaaaagttgttagataggattctaaaa  
cataattcataaataaaaagtagactaatgaaactatataaaaatttaaaacttgtaggaaaaaacatccaaaacataa  
aaagccacagactagagaaaaactcgcacatcatcatatatatttgataaggagtttatatccagaatcatgaggaactc  
ttacaactgaataagatgacaattccagttaaaaacaggcaagatttgaaaagacatttcatgaaagaagatattacag  
tggccaataagcacattaaaaacgacttaacatcattcgttatttgggaatggagatttttaaaaccacagtcagatacagc  
cctaaagtgattagaatccaacgctaatgccatggctttttagaagacagtggttaattctcatgtctgtctctgcattcag  
tctgttgcagtcacatcttttgggttaaacacatgaaaaaacctggccttaccaggtcagtcagttggaaaagggtatagt  
gatcccttttaatatagcctttcagataaattatggaacttacttgatatattatgtctgaaactggacaactggctcattctt  
taaagcgggttgatgggttaaccctaaatgataactgcccacagacttgatttctttagggaacctgggtacagct  
gagttaaagtgtcttctcgttacttaatttataagaaaagcagcctgtatctctaaagacctgttctattttgtgtgtg  
tttagtttttaaatatgcatttcttcttccatagatactcttcttattgactatcagttcaaccttttagtctatcccag  
gaagatgatcgctattacacagctatcaattttgtggctactcctgaacgaagtaagattttttaaagcttctcattttg  
ttttgaatttgtatggatcttttcttgggtcattacggatggacgtactgccttaacagtgctctccagactggagta  
cgagatgatctctagaggtataggaaagaaatgttagactctacgttatctccttccacataaaaaggcaaaagtgatgt  
taataagatttaccaggtatcttagacacagactgacattgatccacgcatacttactctgcctgtcagcccatcatggc  
ctcacaacagaaaagggaactccaccatcagagggcagatagcagagcctgggttactctctcaagagtgaccagaggctt  
aaagacactgctgattgaaatgccagtgatgcagcccaatcagacagcaaggaggggaccccaatcagacagcaaggg  
aggggaccccaatcagacagcaaggaggaggacactctgctcctagagtgagttcttagcctcattaggaggcaaaacag  
caaaggcttagttgggtccattagaagtttagccaggaataatttaaatgtttaaataatgtatgtataaaatgtgggtt  
tttatctgtctcattaaagatgaagcccaactgccttaaggtattacctagtggtagaagggaaggccacactgoggg  
cattttaaactgaaaacatacagaacaagatgcacctgtacagtttcttcaatgaaatataaagtcatgcagtaacca  
cttcagttatttaagaaattttggtaaacataatggtaaatattttaggaacttcttggagatttcttacttctcatga  
acatacacaagccatttaccattacaaaattccattaaacataaatgtgacaaataatcatggaaaacaatatggcag  
taagacaaaagtatatgtcttggtaacaaatgttatctataaacattgctaaatttaatttttaaagttagaacaagct  
atagtgtgggtatgtttactatacaagtaaatgaaaataggatttgtttttaaataattctatttcaaagataaaaatga  
aaatgtactagaaaagatacattctcaaaagtaaatgtctttttaaagtaaaaataataaattactttaaatgaattatt  
ttatggaaaaactataggtattaatatatcatttgagtggtgtattgactggaatttttttcttaacgattttagaata  
agattttaacaaaagtagccatatatgaaatgtattcactgcctcataagcaagcgtttgaagtgtcctaaatcttccagg  
atatacttcatgccattaatgtcattaaaaaaataaatatagtagaatctttgtaataacttcttaaccaggtgggaagca  
accatgaaagtatttggaccttctggatttccagtttatcttacgacagaatacttactagagttatccaaatgcata  
tgttctgtcctctataaaaagcacaagcattttaagtttatgtattcttctgtggaaagacatatagttgaccttgagc  
aatacaggtttaaactgtgaggatccacttacacgtagacttttttcaaccaagcgtggaatgaaaatacagttattgca  
agatgtgaaaacctgagtagacagagggtggacttttcatatgcaagggttctatgggcagactgoggaactggagttatgtg  
tggatttgggcagctcagggggctcctggaaccgatgacccagctataccaaggatgactgtaattgttagttgtgtgc  
tgccagcaataactaaatacaaaataagtaaaacacttgagctgtccttcaagatgaagggtgaggtcttatcagtaaatga  
gaaggtaaatgctttgtgagagaaaacttctgggttaatgatcacattttaaaaaatagctgtttggaaatgttccattgt  
tgtgattttgtgctattaaaatgatcaaaacaacaccttaaaaactcttattctaacctctcaagatcttttaaaaatga  
ataatttcagtacagtcggatgcattctgtaaaagataaaaataaacattgatttagtttgcaaaaataattgtttgaccc  
cagttaaagagatgtactagtcgaatttccagtttgactaattattaatgttataatttacctaacatcccaacagttacac

FIG. 7

14/157

ttcctccactggcttaacagattcctcagcaatatctttattagtcattaaagccaaggatcaaaataaattaagttaga  
ttagccctgtgaactgctatatctcctaagtggttagacaagggttttcaaaactaagaagccatactcaatcatatttctct  
tgaaataatatttttagtaagagaaaaatatttttcaaactcatgaatataataaaattattttttaaattaagttacatt  
cagctctatatgccttttttaaggctgtttccaggttttgaggagaagtttacactgtatataatctatgagtttagaatt  
atatgggtttcagttataaataaatataattttgggcacttgatcaagtggttttattggtaggagtgaaactcagacatt  
tgaaaacagctgccagagaatccttaggaggggtgattccccagcacaggtcaggagatgcaaatcactgtgctctgata  
gggtcttttttaaaaggcactttctcatgcagttaggcatttgataaagcattaaattatgtatactttaatgggagggga  
ataaaattttattggcatatatgtctgttattaaggaaaattgggttagaaaattttgctaactctctgtgagtttctcta  
aagacatcataaaatgtttgtgattaaagaacatttagaggagtaaactttattgctttatttttaaaatctagaaattgtt  
ttaattaaattttctaataatttgtaccgctcatcagcaaaacagggtattggacatgttcatcaatgctccaagaattt  
caacctcaacatcacctgggctgcccagtttctcaggttaaagacataccttagagaagaccctgcaaatgaagggtggtag  
attaaagaaatgtaatataggaattgagaaagcagctccttaggagacagattgggttgaaaaccacccctgcccactactag  
ctatgagaccttgagcaagtatctaaatccctctctaaaccttagcattattttattcatctgtaaagtgaggataatga  
tacttacctcttagaattgttgtatagattaaattagggtatacctaccagagcttgctgtgggtgtcaggtcagtggtgt  
gggtactaccctagcccaccaccacatttctgttcttctgtgtggccactggcactaccatcattgtctacatccgtgct  
tcgggaagtgaaaaaattcaaatgattcgtttcaataaatgaaaacattttaataaaaatgagatttttagtaggtacagaga  
aatgtaacttggggaattacatcaagctctaaaagcacagctcttctgtctgccttactgtgattcactggaagatttct  
gtatagaaaaatctaaagaaataaaggatgaaggccaagtgctgtggctcatgcctttaatctcaacactttgggagactg  
aggcaggaggattgcttgagctcaggagtttaagaccagcctgggcaacatagtgagaccctgtctcaaaaaaaaaaaaa  
aaaaaaaaaccagggtgtagtgactcatcctgtaatgccaaactactcaggaggctgagatgagaggatcacttgagcccag  
gggttgaggattgcaactgagccatgatcgaccattgtactccagtggtgggcaaacagagcaagatcccatctctttaaa  
aaaaaaaaaaaaaaaaaagaaaaaggataatcactacttaacttgataactcaacaagtagatatgggttgaaatttgtc  
cattaaatttacttgcaccgtgctgttaggcaagttacttaaggtttctgagccagtttctcctgtataaagtaggata  
gtaaaaacaccgtcctggcaggggctgatagctcatgcctgtaatcccagcactgtgggaagccaaggtgggaagatcac  
ctgaggtcaagagttttgagaccagcctggccaacatgggtgaaaccctgtctctactaaaaatcaaaaaatcagccaggc  
gtggaggcacgtgcctgtaatcccagctactcaggaggctgaggtaggagaatcgcttgaaacctgaaggtagaggttgc  
tgtgagccgagatcacgccactgcactccagcttgggtgacagagtgaaactccatctcaaaaaaaaaaaaaataaaaaac  
accttcccagtagagtgatgtgagaattaaatagatataaaatgaagtactcaatatagtgcttgaaatgtggtaaat  
ggtaactatattttatcatattactattacaactactgggtttttaaaatcaaaaaacacaaagcaatgagattgatgca  
aaataagaatattgccttgtgcacgccacttacgtttatcatcttaaaacatttgtgtagaatttgagaaaagttcagaaa  
ctctcaatgaggaggaggttttaagaaaaagtctgaattatcagagttatttgagaaaaggcaacatctccaggcatgtga  
aagatttgcaatgagccggggcggtggctcatgcctgtaatcctagcactttgggaggctggggcagggtggattacctgag  
gtcaggagttcaagaccagcctgaccaacatgggtgaaatccgtccctactaaaaatacaaaactagctgggtgtgggtgg  
ggcgtgcctgtaatcccagctactcaggaggctgaggtaggagaatcgcttgaaacctgggaggcgaggttgcaagtgaac  
caagattgcatcattgcactccagcctgggcaacaaaaagcgaattccatctcaaaaaaaaaaaaaaaaaaaaaaagatt  
tgccatgagtgctcatgaagacgtgataatgtgggtcttagtcacagggtctaaactcagacatggaaaaaagttccatt  
tcattaatctttatcggcacttgaaatcctggctaaggagagaattgtggaacattgaaggactctctgggaataggatgga  
gttataccagattagggggacttaaatactgtggtagctgggtgagaaggaggagactgagtgaccccttgaaacctcc  
tcctgtctacagtgggttaggcagtgagcggtagatcagcattactggcatgggagctctggcgcatggccaggaggtgt  
aaaggggaaatgcaaaaggaattgaagtgggtgtgggcaaaagttaattgccagtgcttggttaataaggattctagtggtatctg  
tattttcatgatcatgtgtgtcacctgtttgggggtggggcaagggtggaaggaggtacatggattcctgggtaaaacca  
tttttctttctttcttttttttttttttgagacggaatttcgctcttattgctaggctggagtgcaatggcgcaatcttg  
gctcactgcaacctctgcctcccaggttcaagcgattctcctgcctcagcctgctgagtagctgggattacaggcatgcg  
ccaccacgcctggctaattttgtatttttagtagagacgggggtttctccatgtttgtcagggtgggtctcaaaactccgacc  
tcagggtgagccaccgcctcggtctccaaagtgcctgggattacagatgtgagccaccgcaaccggccagtaaaaccatt  
ttgggttaggggcataggctgtatatcagcctgcccagctttaaatcttatctccatttcttgttggctgtgtgacttga  
gggaagttacttaatttctgtgaacctcaatttccagctctataaatgaagataataatagggcagctgtgaggattaaa  
tgagattgagctcttaaggctattactggaacacaggaaatgtttgataaatgctattgtccttattattaatgaggcc  
agattctgttctccccctagcccccaaaaaatgtctcttctctttcatgtttttcttttaacagctggaaccaggctg  
gagaagagatgcctgttgtttcaaaaaccaacatgaaggagtacaaagatagtttctctaatgagaagtttgattttcgc  
aaccacccaaatatacatttctttgtttatgtcagtaatttccctggcccatcaaaattcaggtaagaagagggttttg  
gtctcatacctgcaaaagggtggtaaatctcttttagtaagtaacaaatttactaatttggagcttgtggtaaatgagatgtg  
caatgtggctttgcctttgtaacgtgtatggcagaggagtgctgagcacatgcagctgcacagaagattggagtgggga

FIG. 7

15/157

tggactgtatcactcatgaaagacatttgcaaaagcactgttgaaagcaagttggcatgtaacagatttgggtcattataa  
acgtattcacttcttcagtgagcatttgccatgtgaaagcctgtagggtacacaaagaacttcaattctagagtaaggt  
ggatgtaagtgaacaaacccacatataacaactgaaagccagagtggtgaaagtgacatgagatgatcccagaaatgct  
atcaaagtttaaaggaggacaaatggggagactatggtgaagaacatcagccttctagtcagacagaggtggattgattc  
ctgggtcctatacaaatcatacagcctttccaagtcctcagttctctgtgcatgtgacctgaagggtagttgttaaggggc  
tgtgcagctgctgcagtgcttgttagcctgctcctttcctctgttcccagggggccagtgactcctcttgtccgaga  
cccatggccccattttaactttttatactcatgtccccggggcctttcctcaataccttctgttcttaccttcttcat  
ttaggtgaatgtggaggttagggataggtgggctttcaaggactgggtcacctttaacatggaagcatgggtcactggac  
ggaggtgttgtgtttgccaatgttcagaagcataatcaacctcagaagcaagt caccacaaacat atgaaaaaagttc  
aacatcactgatcattagagaaatgcaaatcagaaccacagtgagat accatctcacaccagtcagaatgtctgttatta  
aaaagtcaaaaaagaacagatgtggcaaggctgtggagaaacaggaacgctttacactgttgggtgggagtgtaaat  
gttcaaccattgtggaagacagtcctggcaattcctcaaagacactagaggcagaaataccatttgaccgccatccatt  
actgggtgtatatccaaaagaatataaatcattctgttaacaaagatacatgcacatgtatgttcattgcagcactattca  
cagtagcaaaagacatagaatcaacctaaatgccccatcagtgatagactggataaagaaaatgtgggtacatatcaccatg  
gaatactatgcagccataaaaaaggaatgagatcatgtcctttgcagggacatggatggagctggaggccattatcctcag  
caaactaatgcaggaacggaaaaccaggtaccacatgttctcactcgtaagtgggagccgagaacacatggacacatgggt  
ggggaacaacacacactggggcctattggagagtgatggggaggaaggagagcatcaggaagaatagctagtggatgctg  
ggcttaatacctaggtgatgagatgatgtgtgtagcaaacactgtgggcacacgtttacctatgtaacaaacccatatacat  
cctgcacatgtacccctgaacttaaaaatttaacaataacaaaaaaagcaagttatcactcatcaattaggatgcttggg  
actgtgactaaagaacagttgatctttccattctggaaatatggaggacaaaaactatgttgtagtttttctcaccaccc  
tctctcctttttcctgtctccatgggtctaagatttgttaatcctctcatcaggttctcctttgcccctcgcatactcct  
gtgtccctcctcagccagtcctgtgaagcatcagccacgcttcttattcctgttttctctgtctagtacacatctgtc  
acaatggctcggcctcttccctcaccacatcctgaaactgtcctgtcatagtgggcccagtgatggaaattttcagttc  
cctttttatctcattgcatggctttgaatcatcatctttctgtctcctggatttttctttcctcgttttctggccag  
tcttttttaggaaatcctcttctccttcttaaacactgggcttgtctaggattcagacctggctcctttctctctcact  
ctgtttcctgcatgtgggacttgtgtgcccacatctattctgggtattctcagtgctgtctccagaccggcactggcagt  
tgctgctggacagccctacctgggtagtcaaaagctccccgcttcgggtgcatattccttcttgtgatcttagccccaga  
cttccatgcttcttctctgttaaacagcaccaccatcaccctgttgtaacaagctagggctgatgatcatttgcattcc  
tgggataaaaccaacttgagcatgacgtattatttcttactttacacgttgttggttgcagtttgctaatatttaggtt  
agaattttgtatctatttcatgagtaagaatgtgctgtgataaacatttcccttctgtccttgttatgggtttgatcatgtt  
aaactaaacttttaataattataatttgaggagtttctcctctctctctctggagatgtaagattagagttaaactgaa  
cctcaagtaacttggtagctctctcctgtaaaatcatttgatcctagtgtcttatttggggagatacttttagctgctga  
ttcagcttcttttaatagttataggatattttgaatttctcctttatttgattcactttttaataatttttctaggtttt  
atatcatgtttttaaatgtatttttaattttacttttaaatgatacataaattgtacgtatttatagggtacagat  
attttgatatacatgatacaatgtataatgatcaaaagtaatttaggatccatcaactcaaacatttatcatttct  
ttgtgttgagtacatttcacatcttctagctattttgaaatgaataagttatttgctaactatagtcataatgactat  
tgtgccattgaacactagaacttattccttgtaactctatttttataccattaacctctatctccccagccccccagca  
gctggtaaccaccattctgcactctacctccatgagatcagtttttttagctcccacatctgagtgagaaaacatata  
tctttctgtgcccgtggcttgtttcacttaacataatgacctccagttccattcatgttgcagaaaggacaggattccatt  
ctttttgtgactgaataatatttcatgtataatatatacaaacattttatttggtccattcatttggtagggacacgtag  
tttctttcttttgatataataccagcagcgggattgtggatcatgtgggtggatctattatgagcagctctcactactgtc  
ttccatagtggctgtaactaaataatttacattcccaccagcagtgaaactagtatcgtcttctctgtatcctcgccagca  
tctgttattttgtctttttaataatagccattttaactgggatgagatgatttctcattatgggttttgatttgcatttgc  
ctgatgggttactgatgttgagatttttttcatatgcccgttggccatttgtatgtcttcttggacaaatctctattcag  
atcatttggccatttttaaatcaagttttttcctatttgagttgttgaattgctgggtatattctgggtataaatccct  
gttggatggatagtttgaacataatttctctcactctataagctgtctcttctcactctctgtttcctttgtttgtagag  
ctttttggcttgatataatccattttgtogatttttgctttgttggctgtgatttcgacatcttacacaaaacatcttt  
gcccagaccaatgtcctgaaggattttcccaatattttctctagtagttttatgggttttaggtcttatacttaagcctt  
taatccattttaatttgatttttgatgtgtgtgagagagaggggctagtggtatttttctgcacatggatatacagttt  
tcccagcaccatttatgaagaacctgtcctttccccactgaatattcttgactccattattgaaaaccagttggccgt  
gaatatgtggatttatctgagttctttattttgttccactggctgtgtatctgtttttatgctgggtatctgtgtt  
gtgggtactatagctttgtagcatattttgaagtaagggtgatgatgacctccagttttgttcttttctcagaaatgc

FIG. 7

16/157

[illegible]

FIG. 7

17/157

gagatggccatacttgttgccaatgcgagtttttttccctccagttccctattgctccacccctctgaaagtcactgaag  
tgggttttagagctttcattctagaggagctcttccccacacttacttctaagaagtaataatttttaagtttcataatca  
aggagatgttgattttcttctcctgggtttctgggtgctccttatctttctggatcactaacttaaatgaagatccacttat  
gaatgctgcattcaggagcacttctgaaaaataaacgatggcatgccaccatttagtttgccaagctttccttgaaaaagc  
caggtaagttctatttgcaaacgtgtctttttgctgtttggtagcaaatgactgcactattcactgtgactaggcccggtgtc  
accgtgtgttctcagtcctcgcatgagagcctgggtgacactgggtctctctgctttggggaccgtatctggctgtgcgc  
tatgtggagggtgagctcgtgtcttctgggttaaatatcatgttctcaacaagggtgcatgtttctttattgctgt  
tttttccccagagacttagcttttctaggtgtgtagcatggctcttaaatacacacatttagcaactgtgcaaatcag  
aaaaatggagagacttatttaaggaagacagtagggaaatgtctaaatagtgcatgtagggcattcagttaaatttcag  
tatggctttcgatgcaggcaaacgcttttgcctttgattctcattttttaaattagattttaaaaattgagattctcct  
tttttaaatgagattttaaaaattgagattctccttttttaaatgagattttaaaaattgagattctccttttttaaatg  
agattttaaaaattgagattctccttttttaaatgagattttaaaaattgagattctccttttttaaatgagattttaaa  
aattgagattctcatttttttaagcaaatgaactgaggaaatctaatctctacttggaatttgtgtaagcattgggg  
ctttaaaaatggtaataactaaactcaagttccttcaactgaaccaatccaggattcttttccataattaggaacagat  
caaacctataactgctgagggaaatttgaaactgactataaaaaatctttagatgtagctaaaaacaaacaaaccca  
cagactgaaaagtcctgtttctgggagggtggcggtgtgttggggcggtgtgtgctgtttttgatgggtgttctctg  
tgggtgagcacccttccctctgcccctgctttctgaaagctatcctagtgttctcctccagtaagccggggctcagccg  
cttcagctatgtccatagtttggggcggtggcggttaggtattcaagatagaagatgggtggacagaaagtggtgaggt  
gggaagggtggagtggtgaggtctctggctcgttcccgactctgcaacagaggttgttactgtaattgtggagctaatataga  
agtctgctgtgtgggtgtttcatctactgttttatgtgaagggtcagtggttcacaagtaaaaggcaacacagaattgggtat  
tgagaactgggcattcatttagatttttagagttctgtttatgggtgctaacaggatcgctatactgagcaagggtatgtgt  
gaaacttttgttcaactggatgtgtcttttttcttataacagctttatttgagacatcatttacataccatacagttcgt  
cattttaagtagacacaaattcagtggtttttagtgatttcaacagaagcgtgcagccaccgtcactacagccaatttttagaa  
tattttcaccacttccagaagaagccctagcagtcagtcctcattccccctcttccggcaacccccagtcgtatttct  
ctgtctctgtggatttgcctattctggacattttatataaattgagtcacatgatatgctctcacttagcatactgtttt  
caagggttcccccgtgttagtatgtttcagtagcttattccttctgcatagctgaataatcttcccttatatgtatataccac  
gttttggtttatccatttttctatcgatggatgcttgggtgttttcaactctttggctattatgaatagtgtagatatac  
ttttaatgtttatattttattgtataataacagtgcatatccgtacctcctgggtgttaattgtctgtatcttaagagttt  
aagaatttataggttgtgtgttaattctagcacttagggagggccgagggcgagatcgctttagctcaggagtttgaga  
ccaggctgggcaacatagtataaacccctgtctactaaagtacaaaaaatttagctggggtgtggaagtgttctgtagt  
cccagctacttgcaaggctgaggcacaagaatcacttgaacccaggaggcagatgttgcagcgagccaagattgcaccac  
tgtactgcagcctgagcgacagagcaagactcctgtctccaaaaaataaatttttctttatgatgaacccacaag  
gctgtacaagaaccttaaatatagtttgaaacaaagaaagacacatgcttttaagtacagtttaattcttttctttctc  
ttatttctatttttttttttttggagacagagctcactctgtctgcgccaggctggagtacagtggtgcaatcttggctg  
ctgcaacatctgcctcctgggttcaagtgtattctcgtgctcagccttgcgagtagctgggactacaggaacacttccac  
acgcccagctaatttttgtatttttagtagagacaggggtttaccatgttgcctccatcattctcagcaaaactatcgcaa  
ggacaaaaaactgaacaccgcttgttctcattcataggtgcaaacgaacatgaggacgcatggacacaggaaggggaa  
catcacacaccagggtcctgttgtgggtggggggaaggggaagggatagcattaggggatatacctaattgttaaatgacg  
agtttaattgggtgcagcacaccaacatggcatatgtatacatatgtataaacacctgcatgttgtgcacatgtaccctaaaa  
cttaaaagtataattaaaaaagaaaaaagaaaaatgtgttattcctttgtctcgaaagtcctcctgttttcaagcttttgc  
ttgggtgtctcttgtcagctcttttaagtctattgttcccttgggtttcatctggactcttttccctgggcaatctcatc  
cacaccatctgggaacattgcctttggctggacaggttaaatcttgtctctagcagacttctctcctgtgatctccagag  
ccaggaacccactagacttctttaaatttcaattcagccagcagacttcttggcccagagcagcagctgtgagtgagcagc  
aggcacctcgttagaagcttgggtgttggaggtgtgcccatgcattcctgggagccatgagaggggaggggagtgccctgt  
gtcacctgagttcaagtttaagagagatcacaccagagctgatcctaaggagcatggggggcgggggaaggggggtgttctg  
ggtggggaaaggaatgaccagccctttcaggaactaccaggagtttagcagagctgggtgacagtgtatgagcaagaag  
gtggagagctgggagggccattatgaagagactggagattgcaggctgtgaacttgagtcctgatgtgactgggaattg  
ttgagatgttttagggcgaaagtcacccagcatcagtggttagaaagacccactggcagcagttgggagtggtgattg  
gaaggagccaggcctaagtctgggaggacttaggaggtccaggagtcagggtggaattcttctgttagaaa  
cccagggaaggggaccatctgatgacacacatcccatagcaacttcttctcctcctctgcacctccacagcagagtgct  
gtttccactatacgaagatttcttaattggcaaggatcgatttttatccctcagtggtcctcagtgctatgtgctaatacag  
ttggagacactccctaaatgtatgocagatgagtatagttgttagtaagggtttacgggttaggtatcagaactttt  
tcttttttacaactttataaactttataaaggagccgacaccttttctgatattgtcagaacaattgtctgaccagtg

FIG. 7

18/157

tatgggctggctgaggttggcaaccaagaatctgttatctcagctctacacagcgtggagttagaagattggattgttgtt  
ggtagaattttgtccagagagtttagaagatagaagaaaaggagggtgaggacataaacatccaaatttacaaagaa  
gaaagtccagtgctcctgtctatatgtgaatagtttaagagggaaatgattgggtctggggaatgcttcaaaataatatga  
tcatagggaagtaagtgtagtgtggatgaaacaaggttggccttgagctgataatttactaaagcaggatgatgagcgt  
gtgaggttcattatagtgctctttctacttttataaaagcttaagattttccaaaatgaaaaagaagctgagcatggtag  
ctcagctctgtaatcccagcactttgagagcctgagacggaaggatcacttgaagccaggagttcgagaccagcctgggc  
aataaaggtatgagtatatgatttggcctcctacagacagctacaggagcatgcctctgctgggcccggcctggagcatag  
tgcacccctggagctgaagtctctgttcgctttacaggaagtaacagcattgtacgggggtttaccagttaacagtgga  
cagctctctagaacactcaggggtcagggggcaaaactccctctccctcaaaccagcccacctgtcttggtgagaatggaaa  
acggagcttcccagcaagcctcctgccttcacccagcgcacaggtgcacaggggtctgagggctcctccctgtggggcc  
tgggaagccgccccttcagtgggctatggctgcattgtctctgatcttgtagctgtgtgtgggttaacaggcacatcgtct  
ccccatataccagagagtcagtgccacttctctatctgtgtggcactcttccatcactgtcaccctcaggataatggcac  
aggaaactaacagagggcactaagagcagtgatatatactggaggagggggggggctgatgacaaaatcaccttctgttt  
cagtaagaggttgggttaaagtgtatgcctaaaatggatagatcaggaaataacaataaaaacattaatgataaaaaggcaa  
ctgccagaagaactgaaaagtaacctagttcaaggtgggggtccctggagagaagccttcagaagagcagaggaagagca  
gggaacatttgcttttcatgactagcctttcagtgcccttttgatatttcagctaagtatgtagtctctaatgatctctaa  
attgttaatgagtgatagaataggaattctggagaatggtaaaaaggaattgctaggagaaaattagggaagaaagc  
acagagcagttatgaggtgggaggttgggagcagacagatgggtgagtgctctgttgcagggcaggaatgagtggtgaat  
ggcattaactggccatggggctctgaccagagcttgggtatacagtttctgtgcccagcacaaaccagcattaggtgcgtg  
gtgtgtgagggccgggtgatggctcatatggatctgggtccagagctgtgggtctgacatagagttccatgtatgaactttgg  
tgggtgggtgggttttgtttcgtttttaaagaaaaaagaagagttgttttatgtcacagccaaactcaaggaggtgaca  
gcacaagaaccaggttagaattgacttagccaggggtcacggaatctccatcctttgcaaagccaagaagtagagggc  
ctcaaggagaatagaggagcacatgggctcctctctgggctgagctctcacccttgggtggggccgctgggtcccaggggtc  
gttgggagcctgcggagatttctgccagttacatacattcagtttctcggagcctgacaggtgctgccttgaagcg  
gaaagacttcacctcacaggggtggccagctgtcctttctcttctctccctgtcctgtgtgctcatgataatggatagat  
agacatctccatggtagaaataatgcagcccttctgcagtcctatgggtccacagcatatggtcagtagacttagatgag  
tgctacataaagtgttaattagttctgcaccacgggtgtcttaccttctgaaaaggaaaactctgggtattgacttagaaatgc  
cagttttatccttccgaccagaacgcaagtcctgaatgagggccagaggccaatgaaaggtcctttctctctcttttcaa  
atctgtctgattaaacagtgacttttcttagtcattcagtatcattcattcattcatttattcaagtttttgagtagcct  
tcttggtagggcacatagtagtgcaaaaatacaatcattttttaaagtgaagcattttgttatgtggtagtggaata  
taagaaatgccatgcttccatatttttatggtaattctggactcaggttctaccaccatacttaattagaatttcattc  
tcagcctctgttttggcagagaccattctcttccctcaggaacatacagaaacctcagggctgaactgagctcagaggt  
tggggaaggtcccctgggtctgatcttcagtcgctgactgctgcgcgcgattccctggcccctcactcttgagggcagcat  
cctcactctgtgtggcagggcttggacatagaatgacacttaagcaccagagcctcctagcctgccacaaggacacc  
ctacattccagggcactttcagaatgcaaaacctccctccattccctcccttccctgcctccaggttggggccctgcc  
tctccttttgggttcaaagcactgtcctctcttctctaccccttatccttcccatctcctccactccccaaaagagcc  
aaactaattatggggattttaaccttccaattaaaggcgaggattttttttttcaatgcagttcctgaaatgctaaagaa  
attgtccagggaaaggagcacaggaggagttttgttttgggtttgttttgggtttgttttgggtttgttttga  
gacagctctcgtctgtcaccagggctggagtgagtggtgcgatcttggttacgcgaacctctccactcctgggttg  
agcaattctgtgctcagcccccaagtagctgggatcacaggtgcacaccaccagccagctaatttttgtattttta  
gtagaaatgggggttttggcatgttggccaggctgggttgcacctcctggcctcaagtgatccaccacctcagcctcca  
aagtagctgggattaggcgtgagccactgcacccagcctgatgtttaataatttcttacttcagagtttctgaagtgtg  
gctatgggagactagcatcagaatcactggaggagcttaagatataagagacagggatagggcctgagattctgactttg  
ttgttaagctatccagagtttgagagccatggcctagttcttggaaatacaaaagtgtatctgctctgtatgggaggtagg  
aagttaccttttcttctctctcttctacaacagtaattttaaagttaggaattgggtatcttcagacagagaagaatgcagg  
aatcctcttctcaggtatgtctgtctttacacagccatggatgatgtgtagaccaagtttgtccaaccttgcccggggc  
tacatgcagccaggaagggtttgaatgcagcccaacacaaatcataaactttcttaaaacatgatttttttattttatt  
tttttttaagctcatcagcactatcgttagtggttagtggttagtggttttatgtgtggccaaagacaattcttccagtgt  
ggcctagggaagccaaaagattggacacccctatatagaccataattatcacacgcccctgagaggggaagcagatcaatgg  
tgagggagcgttagaggagagcctgccagagtttccagtgggggccggggatgcttaaaaacggagggcagtgatgtgctgag  
ccagaaaggacagggaggacccccatgcacacaggcagggaggctctctggtgacatctgcctgaggtacagtagtaca  
gtagtagagtgagcgggaagcacttttgggtgaagtgcacatttagtgattgaggaagtgcaggtccgtaatatcttacc

FIG. 7



19/157

caaaccttggggctcactgagcttccgagttcagagtttttcagaattgatgaacacatccagaaggcccatggggcaga  
tacaccttagacatattgatgaaagtggagttggaattgggttggtggaagtctgtaggacgtgaattctaggtttcttggg  
gtgtggagtatgggtgtctactgtttggaaacagagaagaatgacattatgagttttatttataaaataagaatttcttct  
ttgtattaaaatagatttttatataggaaattttcattaccttcagtatcagaaaggcccccataatctatttcacogttt  
ccataataaattgtgaaagaaagatttgattgggatattgtctttaaaaatgaaacaaaaagtatgggaaaatacttta  
atlttagagtttaacaacatttatagttacaatcatgaagcaggtatgggtgtctgttcatgagccaagttcataaaagaaa  
aaaaaatcagatcttactgagttctgtgagacagatcattatttttcttattaaacatattcattttattcagcaaatgttt  
attaagtgccaatataatgtatcagtggtgtcttttaggcactggagatatatggtcagtggtataaggggaagtctctgtc  
ttcaatcctcacagatcttattttccataagtaagtgtacaaatcggggtatgaatataatagagcaagataaagagat  
gggggtgtggcttcagttgtgcacatgggtgcccagaaatctgatagctccatcatggagtccttcaacttgagacttcaca  
gcaggatgtaaaaagcaacagccaggaaactccttagttttcccaaagtcctcccttaatctcgttttagagaattatgaa  
acacttccataggcttccatctttatcccttcttcagttgataaactgaagggtctaaatctgccaagcttcccttgccaa  
tgggtttgtgagcaggttctccaacagcgtctccactaaaaatcttgtcttttacaagatggattttacagttgggtggc  
attgtatgacagtcgtgtgagtagggacagcccaatgtatgtttgttaatacagtttatgccttttacattgagtaatttt  
taattacctcgtgtaatttttaacagcactctccttcatcttcagaagtgttcagtttgaacaataagtttgtcatcttg  
tcagtcacaaggttgaccaacaaagactgaccgcagtggtgggggtattcagcaggggttagatgtttctaagtggtcagtg  
ctttcataaacattttcacgttatgacctttgcctcctcacacaatcttgtaactatggagacttagatgataaagattc  
aactaagaggatccctttaagtcctgggtaccagccataccaaaagctgttttatatgaggttctttatgattataatt  
tggacatgaacatggagtggtcctgcttcagcaagaaatcaggcttaataaaccattagcatccagtggtggttaaaggcggg  
ggctgcccctgaggagtgtgtgtgtggcacaggggcgcagtaacaccagataagtcctcagacagtgcccttagcatccagtg  
gtttttggcatctagcagggagttaaagtatctcatagattctttaagtgtaaacaccgaggacacacttaacgttttaa  
aagattgaaatccagacagggctgggggtggactctaactgtttctctctctctctctctctctctctctctctctctct  
cattccccaaaatgacctgtttatataattctccagtccttatctaaaacttttatgtgaggaccatggctcatattccg  
cagctttccttctctcaaggagctaggactgggttatgtgtataaaggaaagaaaagaaaagtggtttctgtttat  
tgtatgaaattgatgtaactttgtattttctgattataaattggtagccagagggaggacacaaaaacttcaaatccagc  
ccataaatatattatggcctgaatagtgttggaaattattttatgtggaatgtcaaaaagaaaatagggccatcaaaaa  
agcaaacagtggaagtctattttctggccctgtgagtcgtggtctcaaaaacagtgcaaggggttgggggttgggtgggac  
taggaaagcagatgctcgtcatttaccttccgtgtttgtcatgagaattgtctgtactgggattgtcaccttggtcagt  
aattgtttcgtttatagtaccattgtctctccaaagaggggtgagttccctaaaaaaaagcaggcagcgaccacaacat  
ctggctagtagcacgtatatctcgtgtatagtaaggacacaattctcgtatttaatacaatttggaggtatgcatttttt  
cgtgggttaataatccctgaaatcatgatgcatttttggtcagtgacttaaaaagcattatgtcatagttaaattgacag  
tgtttttcctaattatacacagaataatgcctttttcagccaatgttatttttagagttgatgaaatagtttgcattttt  
atattcttgggtatgaagactattaactgtataataaaattattttaagttctgatgacgtagtgatgagatgttcagtg  
tagcaaaacgcgagggcacaaattttttgtaaaaaggacaatatcttgagtcagtcacacactgagtaacatgaatgtgacc  
aatgaagagcacagaatggggctgcccacctggaggcagagccatgaattggtactgaggcaaaacaggagagctttgt  
gtgtgggtggaggcgcagggagggagcgcaggcttcagtgaggcatgaggggtctccagggtcaaacagacagaacct  
cagaactggtgtgaagcttagaaggggtgagggaggtgtgtctaaagcctcagtgatttgctatgggactggcaactaa  
acaagttaactaaaggggtgaggtactttttccttaagtgaaaaaagtttcccttaacttctttaagtgtaagctctgttt  
tcagcttaactcttgaaatgagaagcgtgaatccgtggtggtttcatgacctgcataatcacactacctgtttttctgtc  
acagattgccttctctcagcacagcaattttatggacctggtacagttcttctgtgactttcttcaggttaattttgcttat  
acagactctctctgttcttttgatttaggcactagatcccttaataagagtatcagtcctcaaaagataaaaagaataatc  
ctgtgtttccaccacagattaaatttacttcatatgtgttcttgaaaaggttagatgtcagttacagtaataataaactcaa  
totagcatgacaaagtgaaaaacgtttgggtttgtatgttctggcttccagcgggggggtttgcatgtatcctgtattaatt  
ttgcatgcacatctatagataatttgggtcttccatattggggagaatttggactaaaacaacatttaatttaataatgggtac  
tgatgtgagctacacatattgcttatgcttggttcaagagatgattctgggtgattatccaaagttttaagctgaaagtgat  
tttttagtgccctgggattttctttttatttccatagttttttctgtagtcatttttatctttatgaaaatactatagac  
aaataattgaaaaagacaaaaaacgtgctgcagttttttaaacaagaaagcagcaatagcctgccttgccttccccacc  
ccagtccccccaatttaaatgtttttgttttttacttctctgtgtttacacttgggttttcagtttaagttctcttttgat  
gtcctactgtagaagatgaggatatactgtgttacatggcctcccttagcatctccctcacaccacaaaagcatacaaa  
agccccatctctttcatttaataccattgtatcataatgcagtggttaaatcagtgatccctatctatgttatggccataa  
aactcttattcacagctgcattgaacagtgattacactaccacattttttgggttagaaggatcattctttgcttatccta  
aatgtagtatttgcatgttttggctttttaattgtttgatttttctatgtgctaccctgattttatcactatactctt  
gctgatgtacctttagaacagtaaaaaataaaataaaagatcattgtactgtgacagaactgagaaatctcaca

FIG. 7



20/157

tgggtcacacacaggaaggtgctctcttgggttgaatttctctcgaggttgcacacctagaaccttccacctcctgccc  
ccaagtgcacttgttccctgaggcctgctgcacaactgtcatttggaaacttcccttgccttccctctatatgaagtctc  
tgtttcccagatcccatgtcttccacctctcatccattccctcttttactggagcacattctctagatttatgagaaaa  
gatgctgggagataaacttttctgtgagctctcctgactgaaaatataattttatcttctctacaatttgattgaatttct  
gggggcaggtgctatggctcacacctgtatttctagcacttgggagaccaaagtgggtgggtcacttgagctcaggagt  
tcgagaccagcctggacaacatggcaaaacctgtctctacaaaaatacaaaaattagctgggtgaggtggcgcaagcc  
tgtagtcccagctaatacaggaggtgaggcaggagaatggcttgagcctaggaggcagaggttgagctgagcggagattg  
caccactgcactccatcctgggtgacagagtgagaccctgtctcaaaaaaagagagagagagagagagagagagagag  
taggttcaaaatcattttccattagaatttagaagacgttatactatttttctagtgtccagtataatattgagaaggc  
tagtgctgttataattcccactcttgtatgggatccgttttctctcttaaaaagaaaagtataatttaagtgaatataa  
agttacagaaacatttgcaaatcctatgtgtgtagctcagtgaaattacacaaatgaacacttctagagtattgataat  
attatctcttgttccaccagaacacttcaggtctcttatttctctctctgtcagtttcggcaagttgtatatagttttctt  
ctaagtatcccattccatctgttccattctcttgttccactcctgctcagagagagagagagagagagagagagagag  
cttgtgtttccctcactcacaacctcctgctcttccccaaaggtaccatttacctgacttctcaaacatacatgtgca  
tttctatttttgaatgcatatgaattatttcttttttttttttttttttttgaacttagtctctctctgtcaccagg  
ccagagtgcagtggtgtgatcttggctcactgcaagcaaccgcccgtcccggttcaagcaattctcctgctcagcct  
cccaagttagctgggattacaggcacgtgcccaccgcgcatggctaattttgtatttttagcagagacggggtttcaccag  
gttggccaggtggtctcgaactcctgacctcaagtgtaccccgccctctgctcccaagtgctgggattacaagcat  
gaaccaccgctcagtcacaaatgtatataaatgaaatcacacataaaattttaagatgatttgggtgatgaagaaaata  
ctgaaaacttctgggttgcagttaaagttgtgcttcgaggttatagctgtaaatctgtggctctccaccttagctgcacc  
tagaatcacttgaggagcttttgatgtccaagccactcttcagaccatttgactcagaacctctaggggtgcaccccaaa  
catttgaaaactctccaggtgattccaatgtgcagctaaaggttgagaatcactgctttaaagcacatattagaaaagaa  
aaaacattcctctcaataactcagaactgctaattaaacctaaagaagtagaagaaagaaaaataataaagagcaaaa  
attaatgatctaagggaagaagtagtagagaggtacttaaaaactgtcagttggatcttggaaaagagtagaataat  
attgacaaaacctggcgaagattaaccaaggaaaaag  
aagggggcatctcccaaatcccgagacgttttttctccagaagggtatttgatgaacaactttatgtgaatgaatt  
tgaaaatgaacagatggaatgggaccattaaaagaaaaatatgcaacttgccaaaactgacacagaagaagtaaggagc  
ctgaatagttctgtatccatttttaaaacaacactaatttaagttttgtcatagtgaactctaggcccccaggtggct  
cattggtaacataaccttttagggccgggtgcagtggtcagcgcctgtaatccagcactttgggaggccaggtgggtg  
gatcacgaggtcaggagatcgagaccatcctggctaaccaggtgaaaccccatctctactaaaaatacaaaaaattagcc  
aggcgggtggtgggcccctataatcccagctactcgggaggtgagacaggagaatggcgtgaacctgggaggcagagc  
ttgctgtgagccgagatcgcgccactgcactccaacctgggcaacagtgagagactccatctcaaaaaaagaggaga  
atactgtgatcatcttaataatggaggagataaagcatttgataaaatttaataccattcatttttaaaaaatttct  
tagaaaactagaaagagaaacttctggaacactgataatggaacaacataaaaagtcaacagcaaacacttcttaagt  
gtagaatgtagaagtttgagtttgggaacaagacagagatccaacttccacagtttcaattactcaggttcaacctggg  
tccaaaaacatttaagtggaaaattccagaaataaacaatttataagtgcttaaaattgtactccattctgagtagtgatg  
aaatgtcgtgccatcctgcttctgctcccacctgggtaggaatcatccctgggtccagtgatccaagccctatgtctacc  
caccactagtgtcttgggttatcagatctaaaaatcatagtatatgtagggttcagtagctatccatggtttcacgcatcc  
actgggagggtgggaatgttaacccacagattaagcaggacactatattgccaaatccaaaggatactttttgggtctcc  
tcttcttgacttctcggtgatatacaaccaactgactgcttcatgagcagagactctccagcgcctctgacccatgctgc  
cctactgctgcttcttagcctgcttattatgcaatcttttctattcagcctctcactgttgatttttccagaccttga  
ccttggcgcgtcttctcactctgtgttctctgactggactgctgcttccccatctgcagcagacatgccagaagtgaac  
aattccatccttctcaccgggtgttccccacctcaactggcaacatgggtctgtgtggttccaagccacttctcttgc  
tttactcagcaggcactcccttaccaggcccccagagctgtggcgtcagtgcaactctccacctccactccccccccac  
caacctagcatcacagctctcgctggaccttttcaacttgactttggccctccagaatccacaaaaagtcagagaaga  
cctttaaagatcatgtatctagcccttaacttctgtgtatctaaaaataaaattcagcgtcctgggtgaactggcgtc  
cgcttttccatattcagcttgcaactcttctctccacatttctgctaccaggacacatctcctatgtcttcaaaagcaaa  
gctcttccctacactgagcatttgtacatgcttatccatttcttgggaacactcttctcactaccttcttctgtata  
ctttcataccttgcctgatttttccagacaacaacaaatttctccaattctctgacaccaactcagtgctctcacagttcag  
ttccattcagacactaccatagttggagcaaatcctgcaggttaggggtcagtcacacagactgctcccacttcaga  
cactaagcacaaatgggtcccaggtacctcactgctactgggccaactcaaaattcagggcttctgggtgactgctcac  
tccaagatttgacaatttgttagaacagctcacagaactcaggagagcactatacgtagaattacagttttattatcacg  
ggtacaactcaggaaccgccaatggaagacatgaataggacaaggcatgggggtgggtgagcacagctcttcatgcgcg

FIG. 7

21/157

tctctggacaagcctcctccctgcacgtggatatactactatcctggaagctcccccaagcctcatcattctgagtttt  
tttattgaggtttcattacatagggcatgatggatgaaatcactttctattgggtgactgaactcatctccagccctttcc  
ccttcttggaggttgggggatggagctgaaagttctaaccctctaaccacctagttgggttttctggtgaccagcccccg  
catctgaggtctcttaggccccctctcccatgagtcattcattagcacacaaaagacactcctgtccctctgaaaattcc  
aaggggtaagctaaattgtctcccttggttctcactcctgtagaactgtattatcttctcgtgggtattcatcacaat  
ttgaagaatatattcatatgtttacttgcttaatatctgggtcggtagcccaagcttcaggggagcagtgcaactgtctct  
cgttctactcctatacacacacagcttcagcatctagtcacataatctagtaagataattgtgggaagtaaaacttagactct  
tcagggtaagtgcagagtgaggaggaacccataggcaccaacctctcagggaaaaaagcttgtgtcaatggcttgctttta  
agggatgtaggggtctgtattgtagctccagtagacatccttaaattcttggagcatgtgtatggaataagatggccctgc  
tgccaagtacgtacagcctgggtgaggaccggccacttctactggacattcagtcctctggggcagccgaatgacactatcag  
aataagtcgaggaagtacttttcagggcccatggcctcttgtgtagtaaaacaggacagatggcattttctcacttctcc  
cttgtgggggttgaaactctggaaaagaatccggaactttgggttttctagacgatgtttactaaactagcgatgataaac  
tttttagctaagaatagagttactttctggcatggggaagaaaaatatttccagggaagtaaacatgtctaaaagaa  
tttatttagcttt  
tcctagcccactgcagcctcaaacctcctgagctgggtctcaagcagtcctcctcactcagtcctccaagtagagctggga  
ctgcaggcatgcaccatcatgcccagctaataattttcttt  
ttaaacttctggcctcaagcagtcctcccatctcagcctcccaaggtgctgggattaccagcgtgagccactgctccag  
cctggtccaaaagaattttaaattggagcctataaggaagggggaaggtatccagataatcctaagcactctgaagaat  
ggctcatataatatatggaaaaaagaaacaaaagtcgggggggaaatctgatgttcaaataagaaaaatcaatgtaaagtga  
gttagaaaaaatatttccagactccgggacagaaaaagtaaaccttaggtcctgacacagtcgtgtgagactgtgtccta  
atgagatgcttttccagcctcaggaaggagcctccttaaagaggggaataagagactctcctctgcttgcttttctgtag  
cttctttagctggagcagcaagtagaaggagactggggaactgggtgtgtcaacagatcttcaactgaatacagaagggcc  
ctctcacttccctgaggcatagagggaagcaggcacagactgtattacgggtggaccaagagaggtttctgagtgaca  
ccagggtcctcgccctctcagcagcagcggcaggctcttgaccaacttagaagctgttgccgttggtctcatgctcttct  
tgtggatttgcctccttactgtttatctatagaataaaaaataacaaaaatcaggattatgcgtggaagccaagaaag  
actgagaaggttattctcagacctctttactgggtcttctcgttctcctcaacttctcattgttaggggtggccccaggac  
attctcatctatctgtattcactctcttgggtgaatgtggccagggtttattgcttgaaataccatatatccttatgcct  
tccaaatgccatgccagcccagacctggcagccagcctccagggttagaataatctgcccaggacatctccctgggctgt  
ccggtagatatctcagggttctgcatgtctaaaatggatctactgatattgtccagattccaaagtgccttttccacacctct  
gcttccacagctctgggtttgagccaccatcatctcacttagactctgggtctccctgataatgccattgcacctctaattg  
gtttctcagcaaaagttaaatcacacactctgttttaaaacctgcactggctcttcccaatgcctcataatgcttg  
cagtgccccacatgatctgcccctcaccctcacccttctctctgctttgtctctggccttacacaccccactccaaaca  
tgctggcctcctgaggtccccaagccatcccaggcagggccttgccctcgggctttgcactgactgctccttttccaga  
cgacttttgcaggcgtccatgagccacctcccataccttctccaagcttttggttcaaatgtcaccttctgagaagaaga  
ggccacacactctgaccatcctgctgtggagtgccactggtaactcctgaaccccttgctcttagcatctttttattct  
gtagcacttacctaccttttggcatgctctgaatttgccttattacttactgtttgtctccactgctggaatgtaaat  
tccatgggacagtaaatctttgctctgtttatcagtggtttccatgagcctagaacaggacctgacccaattaaagtattg  
attaaaaagacaatgtaggccagtcacgtgaggcctgtaatctttgccaaggcaggcagatcacttgaggttaggagttt  
gagaccagactggccaacgtgggtgaaacctcgtctccactaaaaatacaaaaaattagctgggtgtgatggcgcacacctg  
tagtccaagctactcaggaggctaaaagaggagaatagaatcgcttgaacctgggaggcagaggttgcagtgaactgaga  
tcgtgccactgcactccagcctgggtgacagagcgagactgtgtctcaaaaaaagacaatgtagactgtcagaagagtt  
ggcagtgctcctgatagcagagaggccagagatgaaggctttccaccaagtgcagatcctggggggggttggaaggcaatc  
ctattgaattgcctgttgtcagggcacccattggaggagctgcaccccttgtaccccttagggaggaggcagtgagcctt  
gagataagggcagttcagtcactcctagcatcttagagtcctctctccagccagcccaggagggtgtgtgtgtgtcagc  
agttgggtgcactgaggtgccattcatgaagctgaccccatcctacaaacgcagagctcagccagcaagagagtaacctg  
atgacactggcagagagccttatgtctacatgtagcactctggaaaaataccctgtagctagcccttttacaagttaat  
gagtaaaaaatgcagttatagcagttatctatttttttacttggtaagcttattcttcttttaacccgaagtcaagctgtt  
ctctttgccatcactgtagcttttcaaatcatccacacctgggtttatacatgagagagcattgaaatcctctgcaacta  
ttaaggccactgggagttattttctcttaatatatttttaagaaagttgttggcttgggtttgatcatttcaaatgtgtttt  
attttacttaaaaatatgtatttgaataatatgaaataacacttaataattacacacatttcttctgtttgcctcttcc  
ttatgaaaagttgattgttactttaataattaactttttcaccttcttagattttaccagtaaatatctgctaacttgaa  
taaacagcatgactaatatttggttataattatcttagaattatttttcttatttgggataagccactgtgggcccattctg  
atgaaggccaagtgggaaggtcagggctctctctgaataccacatgggaacgatacagtcagaggcctgggttaggaagcc

FIG. 7

22/157

tggttacccaagttgtctcactcattacotcatgagccggcacattgtgtgacaaagcatcagaagggagcccgagg  
tggeatgtgggagctgttttactgcctgggcctgaggtgtgggcaacggtgggagctgattctgggggctgggataata  
actcagtaactttcctttctcacagttgtttcctctctttgctcctggtggctgctgtgggtttggaagatcaaacaaagt  
gttgggcctccagacgtagagaggttaagcttcagtgggtaagattaaagaatcctggaagagcttttttctctctt  
ttctcttaagcaagtggttttagctatttagtgataatggacagacagacatctccacggaaggaataatgcagccct  
tctgcagtcctaggttctgcagcatatggtcagtagacttagatgagtgctacataaagtgtaattagttctgcaccaca  
gtgtcttacctccgaaaaggaaactctggtattgacttggaatgccagttttatccttcagaccagaacccaagtct  
gactgaggccagagatccaagggcgtattagcacagccaccgcgtgattaggcaccgccttcagtgagaacttgtcatgag  
agagggaaatcagccatttaagcatctcaaaaatttttctatttcaaggaaagataaatgtgtgtgtagtttaagtttt  
aacttgagttttatttttaataagtcttgatttgtttccagagcttcattccaaaagtaataagcaaatatttacagct  
gacaaatttaataaaataaacttggttgcaattacaacatatcaaatgcccatcaatcaatgagtggtataagaaaaatg  
tggggtgtgtatgtgcgtgtgtatgtatatatatatttatatacatttgcagcaacctggatggaattggagaccgtt  
attctaagtgaataaattcaggaatggaaaaccacacttcatatattctcacttacaagtgggaactaagccctgaggat  
gcaaaggcataagaatgatacagtggttttgaggactcgggggaaagagtgggaggggagtgagggataaaatactata  
catcgggtacagtgtagactgctcgggtgataggtgcaccagaatctcagaatcaccatgaaagaacttatcatgtaa  
ccaaacaccacctgttcccccaaaaacctattgaaataataaataaataaataaataaataaataaataaataaataa  
aataaccaccattaaaataattacaacataaaaactttgtgtatttgcaactttcacttgcacatcatctttgcctttcat  
ttagtcacgtcgttggtcattgttggcattgttgcataatggaaatgctcacgtgtatataagatctagacaaagaagctgt  
ctgggggttttttggatctactcatgcctttttctatataaaaatgtatttatagataagatttagaacagaagggaggt  
atgtaaaatcacaaactcacagcactcagcactgttgataggatattcatgtctgaatagggttaagacaggtccagga  
tatgggcctcttgtgtgtgggccacagtagccactttctgacccacataaagatgatgttatcatagagggatgttgca  
cactctttattatttttttaataagtggataacatatcaaaagggttttttagataactgtattactgatatacttt  
aaaagcctaataatgaattacgtgatgttctgaaaaaaatttatatttcaatttaatttaaatgtatatatttatcacaatgga  
gtaagaaagaagtgaagaagccacagtaaacagttatcaagaaaggagttgaaaacttttgagtgtctactctctccaaga  
aggacattagaccagggaggtttcaaggttgagatgagcttatggagattatgatagcctgttctcctcctgagtgcttat  
gagaaataattgtttctaaaatctgtaaagcaaaaggacagatggaaagcttgaaagagtcggaattgattttatacaca  
taaatcttggtagcaaaaacctgacaaaggttagcaccaaaatacaaaagggcagctgcatagtgaatgtatgtgcagcagtc  
taattgcagtgtaatgaaagagtcacttaataccagaaaacctattagaatatatcatgtcaaaaaaagaatatatcatg  
tcaatggggaaaaaataaacacctgaattcttctcagtgggccagaaatcatttcgtcatttatgaaactcaatatccat  
tccataatttgaaaacaaaactcacagtaaaatcaaaaatacctttttaacgctatgaagaatatctggataaccagcatc  
actcatgataaaaacaaaacacacagaggttgatccaggctagttagttttaacaagagctgacatctgagcatttaa  
aaattccaggcactgagttctaagtggttttacgtagatttaactaatccttatctcaaacctatgaaataggatgtgtcc  
ctcgttatgttttatggatgaggggactgaggctcagaaagggtgaagtgaactacaccacatagtttaactaagaggcaga  
cgaggtcttaactagggactctgtcttcacagcccaatctgttatcttcgtgtaatcctacagtcattcagctttgtg  
ctagctgttctagactgtgcactaagtgaaagggcactaagaagtggtgttactatttgaaaggaacaaaattattatt  
ttcatataaagttattaaaacaatgtagtagagatttgctgcataaaaaatgaaggcttaagtatatcaaaaaataaatg  
aataacacagaagctaaacaatcaactggaaattgtttgcagtggtgcagagagaaagtaagatgcacagcctacaggaa  
aaagattaaaatatatatatgtgttagtatggagttgggtcatctaataaactatgcttttaccctgtgatttggaaatt  
tctctgtaactgtgaagctgaccattcagaatatatttcatcatttttaaaatgtcagccattatttattccctctcaga  
attgtagaagcaatgatttgatgttatttttaaaaacaatagtagtagcggctgacgctgttatttattccctctcaga  
agccccgggttttaagtgcttttacacatgttgccctcctgaacctgtcttattgtgacccagcacttaagcacagacc  
caaggagcattgtctgcataatattagtcgagtaaatggacctacagcttttaaaaaatgtaagaatgttggtattttaagaa  
aatatttataaaaatgcccttctagaacagagccatatctatccaggccaagattctgtctctacagtagttccaagga  
ggcagtttgaaagtaaggcctacttgtatccaaggcataaacctcaagcttcagcatcatgccagaaagatctctaata  
tccctaatgatattgtgaacggctgttcacatgtgctagggcgtgtgctgtatctcatttgatcccagtgacaacct  
tatgggctaggaacacatttcagaggaggacacggatgccagagacttcaaacacctttcttaagggtcacacagctgt  
agactattggagcaggagttccatcaagcctccttaaatcccgagccctgtttgtccctctactotgcaaaatttcac  
atttgtttgttttagcttataccacctctctgaaggaaaggttcttccagttcctatctgcccgtataatgaggcatagg  
ttttgtgtgtcctgtgaatccttggtgccagtttcagtaggaagaaagcaagcctgacttttagtagtagaaatacagagga  
gatagggccacacattttatgtatttgacaagtgcagctcatagcattaaatccacagtaactttattttttatctttat  
tttctattttactttaagttcttggtatcatgtacagaacatgtagatttggttacataggtatacatgtgcccgtgtgg  
tttgccgacctatcagcctgttatctaggttttaagccctgcagcttgatttgggtatttgtcctaatgtctcctccct  
tgctcccccacctgacaggccccagtggtgtgtgttcccttccctgtgtccatgtgttctcatggttccactccactt

FIG. 7

23/157

aagagtgaggacatggagatatttgggttttctgttcctgtgttgggttggagactgatggccttcagcttcatcogtgt  
ccctgcaaaggatataaactcattttattttatggccgcagatagattccatgggtgtatgtgcccagcttttctttatcca  
gtctatcattgatgggcatttgggttgggttctaagctcttcgctactgtaaatagtgctgcaataaacatcgtaacgtta  
atgtatctctgacaggggcttttaaaaaatataaggacaatgggtgttatcacatccaacaaaatttagtaatcccta  
ttatctaataatcaagctcaaaaagtatcattttggacttgatttgggtgaatcagggaggatgaggtgttttaacatgt  
agcatttaactaataaaaaccccttcagtaactcattccattaatggccttgccatattatttatcagcaactcttcgagag  
atgcaacagatggccagccgtccctttgcccctctgtaaatgtgccttggaacagatgaggagccctcctgatcttatgg  
ggggagtataaagggtgagaatgtgactcagaagtcctataaacttgactttttaaacttaggctcctaagctcgggaaa  
ccagagagagacaaaagccctattccattacccctcttttttctctctcttttttaacaataagctaaaacccgaagt  
gctgggtactcttttgccttttgttttcaatctttgtttcatgtcgtgagcagctctgtgttcaggggatttttaaaatc  
taaatgatattgtctgcttttcttttaagtattctgtggcatacaaaaacttttatgtgaaaagagctacattgaacatcga  
atcattgttttttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttgttt  
cagctcactgcaatctaccccccgggttcaagcgaattctcttgcctcagcctcctgaggtgtgtgggattaccgatgtgcg  
caaccacgcccagctgatttttgtatttttagtgaggatgggggtttcaccatgtttggccagggctgggtctcgaactcctga  
cttcaagcaatccacccacccctggcctcccaaagtgtgggattacaggtgtgaaccacgcacccagcctatttttctt  
tataatttggggaattttttaagcagcaaaatgataatacactgtgtattaactctgtggaaaatatttatcccccc  
ctccccaaaatgatgaagaaaaatggtaagactttctccttgggcaaaaatgatggaaatatttaagatacgtcgggagaata  
gctatgtatcttgaataaaaacataacttacttaaaaatgttactgagctcacaggaagtggtaaaaatctccaggggaatact  
ctccttccctaccctaaaaagaacaaaccataaaacaaaaaccagaacccatgatttgccttaagtgggtccagatggtaaa  
cacaacgcccacctgaatagatggaaatcctctctatagaaaaacatcttcaattcagcgtgttgaagtcccacagatc  
aagatgaagtaattaaagatcaccaggggacaagaggcaagccgcccataaataagagtgcagaaaaaacagatgattgat  
gtggacaggactccaaatgttagaattgaaatctaaagaatataatataactgtgtatgaaatgtttaaagaaataaaag  
atgaaatcataaagatgagagtcaagtatggccaagcagatctgaagaagaaccaaattggaactctcagaaatgaaaaat  
atacttgttgaattttttaataataataaactcaatttcatgtttaaacagcatatcagacataagagaatttacaaggt  
ggaagatgaatctgaaaaaaagtacgcagaatgcagctcagagagacaaggagatagaaagtacaatagatactaagaata  
tggaggatagaatgagaagatccaactcatatataattccagaatcccagaacaaaggagagggaatattcagaaatgat  
agctttccagaactgatggaaaacatgaatatacagattccagaagcaaaaacatattgtaagcatgattaaagaaagaaa  
gggaaggcaggaaggaaagaaaaaagttggatttctcataatgaaattgtataactccaaagacaaagagaacatctta  
aaatcagccagagaaaaagatcacatacaacgtgggaaaaatgggagtttagtgttagtgcaactgttgatgaaatgata  
ggagccactgtggaaaaacagtaaccactattccccccagaattaaaagtagaattatataatgatctggcaatctgggttt  
tacctagaagaactgaaagcagagtctcaaagacatatttgaacatcagtggttcttggcagcattattcacagttagccaa  
aaagtgggaagcaggcccaagtgtccatggatggataaatggataagcacaacgtgggtctatgcatccaatggaatgttat  
ccagccttccaaaggaagcaaatctgacccatgctacaatgtgatgaacctgaagacataactaagtgaataactccag  
taacaaaaaaaataaacattgtatgattacacttttaagagatacacttgtgttagtcaagctcataagacaggaagttag  
aacagtgggttcacaggggctgggagaaggggaaaaatgggagtttagtgtttaatgggtacgaagttagtatttacaaga  
tgaaaagagtctggagatggatagtaattgatgttgccacaacatgtgaatatacttaataccactgaactgtacattt  
ttaaatgggtcaagatggtaaatgttatgtgtattttatcacatttttgttaaaaatgggaaaagaggccgggtgtggtg  
gctaacacctgtaataccgagcactttgggagggcggaggcggggcgatcacttgaggtcaggagttcaagaccagcctggc  
caacatggtgaaaccccatctctactaaaaatacagaaattagccaggcatgatggcacacatctgtaatccagctact  
tgggaggtcaggcatgagaatcatttgaacctagaaggcagagttcacagtgagctgagattgcaccactgtactccag  
cctgggagcagagaaacagactctgtcgcaaaaaaaaaaaaaaaaaaaaaaagggaagatcacctgcaaggaatg  
acaagctgacccgctaacttctcagaagtgcagaagagagtgagagaatgtcttgaagtactgaaagaaacctccccacct  
agaatagtgtgcccagtaagcaccttccgagaacaaggatgaaataaagatatcttcaaattagaataaaaaattgagaat  
ttgccaccagcagacctgcacaaaagaaatgtctgaaggatgtgctaaaaatgaaggaaaaatccccccaggaggatagac  
taaagtgtaaagaaagaaatagtgagtaataaaagtataatcatatagaaatctaaacaaacattatcctggtaaaaa  
atatttgggttataaaaaatagacaagcctaaggtattgaacaacatgatataatgtcagcagcatatgacacaggttagt  
gctccttaaaacatttcttggattttcttccaggacacacacacttgccctctagctgttcttgctcagtggtctttgc  
tcagtctctctcatcttttgacagtggtgtgctcaggccctgctttgtcccttcttgtgtcttttacacccacttct  
tcagtgtatcttactatcacaggtttaagtaccatgatcctctcacctggactgctcccatgagcctccaacacaaacat  
cagttgcctactcagtgctcccacttggatattttaacatggccatggccaaactgagctttggatgtttatttctcaga  
cttgcctctctgacataccttccgctttcagcacataaacctccatcacccctagttgctcaggccaaaactcgtggaat  
tgttcttgactcattcttgatgtcatgttctcattggacagccagctctatcagcaaatcttgactccgtattcaaaaatg  
tttccagccactgtgtcccatctctgcccactattccacttgatctaggccaacgctgtctctcacctggattcttgca

FIG. 7

24/157

gttaacttcataactgatctccctgectctgccattcacccatctagtcagttccttgtgccacagcagagtggtactgga  
aattatcttctcagccaaactgaagaaactgatgacatttctgtgggtatagtcagcagaattctaaaggtggccctcaag  
attccatcccacgtttattcagtcacacactactctaggtactgctagagggatttgcagatgtaatgaaacaaggga  
gttgactttaagatacagaaagttctctggctgggtggcagaaagaggaagccagagagattggaaacatgagaaggact  
tgacacaccatccctgggttagagatggagtgccacagaagtttttggttttgtgagacagagctctggctgtgtcac  
ccacactggagtgagtgccatgatcatggctcactgtagcctcaaactcttgggtcaagggtcctccctctgggtc  
tcccaaagtgtgggattacaggtgtgagccactgcgtgaagccctgaatgcatttcaaagcagagctctcccccagagctc  
ccaactaagagtcccatccagctgacaacttgacttcagcctgagagacctcaggagatagctgataattccccagaact  
catggaagacatgaacagacctgtctctctcggcagagaaccctacagagcttgctgggacttctgacctctaaagctgt  
gagataataagtggtgtgttttaagcaactaagtttgcagtagtttgttacactaatataagctgggtgcggtaggtc  
acacctttaatcccagcacttgggaggtgaggtgggaggatcgcttgagcccaggagttcaagacctgctggacaac  
atagtgagagacccatttctattctttaaaaaaaaaaaggaaaggaaagaaaaagagaagaaactaatataggggacaaat  
ctcatctaaagggtgcaacagaagatgaaaacataatttccagagccaaatgcacccagagagtagacaaagaatctact  
tggtcagccaaaccagcttttacaagcttttacatccagcaggttctctacttagtgtctaaaggacagacaaaataga  
atgttttttaagttctccatgatctgcatctgtgttgagccaggtatgcagaactatctgacaaatctaccaagcact  
caccactgccatctctgtgacctacagttatgaattagaatttcaaatcgagccaaagtagacatccagaagaaacca  
cttggcgtggcccttccctacagttcaagcagccctctgggaagggtgactgcgagacctcagtggaaggtagcgaagcagt  
accacccgtgtcactgggtgctcagagagtagctgttagctctctcctgcccttttctattatcttaaaaactgcaatgtac  
cttttttatatacctgaaattgatgggtcttataagcacaatgaccgttaagttgtccgtatcattaaagcttgagta  
tcacttaaggtaattttgttttctctgccagactgttcccaaacccattgcactggagccgtgttttggaacaaagcc  
gctgtcctctctgtgtttgtgaggctcctcgaggcctgggtggcatccctcctcctgggcagtcagggtgagtagatgcg  
gtccagcgaagacaccttctaagcatgtgagggagctaagcatgggataacttcccttctcagagaacaaggttataaa  
gggtataaccaacagttcctgcccattgccatggcaaggattcagtgagataacccattttcacctatatgccccaaactg  
tttcccgaggtcactgtccctcctggggggacgtcccggaacctcctcacctgttggcctcccagcctctcagtgcc  
tcgtggatgctctgaccacaccttcccttggctgtcggagtttctgggtccccccttgcctctgtgtctttccctgtct  
cttttgccctgtcctgcccctgaatctgggtgtgtccttccacctgcagttcttctcagttgcctctccactcca  
caacttccagccatctccccctgggtgtttcttcaattgggtatttctagccctcctgcatttctttgtcttttcac  
tgttttaaatataagggcacaaacacacgataaacataaaggaaacgtctgaaaatccacgcacagttctgtctatccaca  
acccaccaattttctattgtttctgtgttctttccagccctggccagcaatttagttgaaatgattgtatagatctaatttt  
tattctgtctttttcatataactatcttaacatttttatagcctactacataaatattttaccaatgcagtgtagacttc  
tccatcaaatatatagaatcatgattatataaacttctatttgaagacagaaaatataggaatacagttttgaaagtatct  
cgcttaattacagagtgctttggaggatgcctcaatgtgatcattttagacaactttcagttgagatgattctggagtc  
gggttaccttagccagagagcaccctgtgtctgggtggaagggtgcccagagatactgggtccagccctgagctccactg  
tcccttcagtgacacctcagacccagaagatcctccagccagctttagaatgggttaggatctatagtagtcttccagagg  
gaggcctttcgacatggctcggccaggattcattgtcttcatctcacctagggattcctgtgagttgagcccttaactcg  
gccttgccctgtgttgaggacaggacaactctctcgtctgtcactcaagtgcagtttgtgtctcctgccaccgcccagccc  
agcaaacagtggtgtgtgtgcccctgagggcaggggtggcggtggggcggaagggtgacaggtgtgcagccctgtagcctt  
tccgtgtccagcctttgctgacatgcccccaaccacctgaaacacctaataaagggtctgtcctcaatgcttgagacccaccc  
aacagtggggcccgtccttgtgccccttctgtacccctgttttaagtggtgcccctttctaggctcccagagcattccc  
cactgtgtgccaggcacttacagctgtctacccccacccccactccccctcacaggctcaacaaggcaagtaaacacgctc  
cactaacacaactcctggcaggtctcagcagagctccgagcctgcgctggccatttcccacgtatataggaatccagct  
ccccagaaccgtgccttccctaagtgatgtcagattactccgttatcaatacaggaaggaaaatgggattgcaagtcctc  
caaagcaaaagtggcgccggcgggcagtgatgagcataaagggtagttgctattgatcgtgggtgtgcgcgccagataact  
aagactttatgcatgccaggtactaagactttgtcagttcactcattgattcttcacactagccctgtactactattatc  
cccatttacaatgaggaaactgaggtgcaggaggttaagaacacctgttatccaaagcctggcccagagcctgc  
cctcctgtatcgtggtatggaatggaagtgcataatgtgactgagtgataggggtgaatctgaatctcgtcataaagc  
gtgctgtctatttctcggtccttttgagggaaggcaggagaaattgggcacttagcacagctctacaatctcatttatat  
gttcccaccaccaagcagagcactgtgtctgggtgatgtgtgtaaaagcaaaactgcaccttttgtgtaaaaagttagggaa  
aggagaagaatgtgtatttgcataagtataccctcaagaatacataaaaaattagtgaaagcgtttaccagat  
ggaggggaaggcgaggagggtcagggtggaagggaagactttccctgtgtatgtgtaccttttgggggtttttgtt  
ctgggtttttgtttgtttgttttgcccatgggattgaaaaataaatgttttaattatttttaaaacaaaaataaaagc  
cagctgtctccagcatatgttctctgtgtgttctcccaaggtctgtgtggccagcgcctgggtggacatttctcagc  
agatgccgatagtgtaacaggagaagtcaggagccgtgagaaaccggaagcagcagcccccctgcacagcctgggacctgc

FIG. 7

25/157

atctgatgctggggccagggaactctcccacgcacgagctagttagtggtggcacaccagagccatctgcagggaagggcggtgg  
 cggggaaatggctgtgctgggtgctgggacggaagactggaaaccctcaaagcatctgactcacctgcatgatcacaagcttt  
 ctttgacgggtttctcccatcctgtgtccagcatctaaccttttacttttgcataaggaaatacttgatttaattacagggtc  
 cagggatgagctgatgggtgctggaggaggccagtgtagagccagtgagagaactaggaatgacactcagggttcaactgtg  
 gaaaactgttcttgggactgtctcaactgtgcaaaaaacaaaagatggagtggttacaagtagacattcgtcatcagttg  
 ttcttgaacatgggtcttttaaaaactagtcagatgaattaaacttggttttcatctgaagcctgctatcttttttaaaagat  
 gtgctattttattcttgcacgatttaggcaattatctctctccaggagtagctttttttctagttgagaatttaataatg  
 gtccatctcttttgatcatatcaagctaggatagaaggggggctattttaaatgtcaagggtcagcagtggttactttgaat  
 gtaaaactgggtataataggtagttttctatagtaacttgatttaatttagtcttaatccatttgaaactctctcttcttct  
 tctctgctgtccctctcttctccatctcaccctccctctctcacacatacacacacaaacacatacacacaacactaa  
 gtgcttagacttttaaatagatctagcaattggaaagtttagtaagcctaagtttttacataattgcattoctacattcttg  
 taaaatttaaatagctaccattggcaatctgctttttttctaaaaatctgatttgagccaggaaagaattttctcaccca  
 aggaacatttgatctagcagcagggtgagaggaaagcagaaatgaatgaactgtgaaagctcctgtttttattatcaaaa  
 aaggacactgtcaagaaggcgccccctgccccacccccctgtgcaccctaggcctgat aagcgatcagaggaaaggactc  
 attcatgtcacgcttcccttgagcagaaaagagcactgagagcacttgggacccctggatcagagagcatctgtgtgtcct  
 gcagcctcctctgaacttgtggttcttctcaggctggggtggactcagatgccaggaaaggacagcctcccatctgtca  
 ggcaagctgccccaaagcctggagaaggacttggttgcctcttctccccaggagggtcgcagcccccacccctccct  
 ctgagaccaaggtgggtggtgctgaggaggggcagcaaatgctgacaaggatgaaaagcacatggaaaaaatggacgagga  
 gggaaaactctgcaaaatggaaaatgaccaaatttaagagggtgggacagtcccctgctcctctcccagaggggcactgct  
 tggaaaattgtgttttccccatttatgggtgctctgtattctggcattatgcagcagcctccagaagctctctctgcttc  
 aaaacctgggatctctggcattaccctattgggatggacccgctggacagcaatgctcgagtttgtgaatttggagagata  
 ctcaaaagagctaaaactgcagcattttaccttttaaatgcagtgccctagagagagagtattgtctctccccaaactaa  
 cccactcccatgaagaattgcctggaaagatgttttcaaggaatttgaaccataaaacactatctgatgcacagaacac  
 ctctactttgagactcaactctcataaagcttctttttcacattactgttaaagaccagacgttctagaaaagacccctc  
 ctctcatgagctccccatccctgctacagaacacagcaccatggcgccctgcagtggaactggcccttaattccccag  
 gccccccagcaaggccaaagggaggccctgggtattgtcctcctacaaggaagatcctcttgtttgttcaaaggacc  
 agtttctcaggccaaagaagtctcttccccatggttagtcctatgccttgaaatatcatgcaccatgacccacagccatc  
 tgggtatgtcttatttttttccctaaaagataatgttttatttttaaaaagggaaggagcaagtgaagtttcttctg  
 tccagcgggtggggaagccgctgaatccactgcttctcctttgcaaccgacagcaaacagcttctccggcctcagggca  
 gaaaaaggggaatggcaggggagtgaaggcgctgggctcggagcctgttccaagaagggaattgggtgtcatctggcagtg  
 ttgctgctcacaagagagccctgtatataaattaaaatagtcaagacaacactgacctgtcacttgtacataactatacag  
 tagtgtccagaatgttcagacattcggagtgtagataaaacagaaaaaatcttcatgtatttttattaaatataacaatg  
 tctgagtttccactaagatgttttgtgcatatgctggatccagggtctcggcaggcccgatgaataaatacaaaa  
 cccaagaacgcgcacccattgtgtgtagtgcctcaggtcctggcaccataggtatttcttaaaaaggagactcatc  
 tgtcagagtgacatgaaaaatcaggcagggaatcgaaacgcagcgcctggaggagactcaggaagcagaggcgtcctg  
 ccgctgcccttggccctgcaagcacatcatgacccttctggcagcctcctgggtgctctgggtagttagggatgaccagt  
 cttgtcctgagaaatgttctcttagtctttaagttcaaagactaacctgtagcaatcagactttccaaaaggggggttct  
 ccattttttgtagttttgtctaaatttttaattgaccatttctcgggaatcagtttattatactgaaaactgggggtgggag  
 tagggagctagtttgttgataaatagttccattttcccggtggagaatttgacataccctggactcctgtgtgctcctg  
 coactccctgcacacagcctgggggagaagcctgtgcctcccctgtggagagaagggaaccccagatccctgagctaac  
 cggaggaaaggcagtcctggacagaagactgtcagcagaaggaaagtactggactaccctgggtgaagtcctgccattca  
 agactggagacacctgggaaataaaaagagcaggggcactgctgggtgggaaggagcattttaccttccagtgcaaatcctg  
 ctccctttgatttaattgggggtgactggggccaggggctgattcacttccctggggagatgggtgggtgttttcatgaacatct  
 ttgatccttccatttcttatttcatccatccattcaacaagtatttgctaaacactaaacttaagctaatgctagggtag  
 tgactgagatgtaaaaatagatttttagaattaaaaacaaaatccaagtcctcacaccctgtcatcccaggagatcttcc  
 ttgtgggtgtttctgtgagaattggccatcctgaggacacagccaggacggcagaggccctcctggcctcaggggcatgcc  
 tgccctaccttctgaaatgtttacccccattgacccaaacttggtccagccattgcgggtgggtttctagatagccaggccac  
 caagagatattggcccttgatgagagtcacacacccctgcctacaaggagatgtttgaaatggagaggaaaattggcacc  
 tcatcttttaaggcagtaattggaattgattttcagtaactgaatttgcacaaaaacattctaaacactagtgaagcct  
 gtttctgtgaactaatlaattctggctctggaaatgtttttgttttatagttatttaagatttctgtttgtttggattcaa  
 gcttagtttgttaatatgtataatttagcatctattacactcatgtaaatatggagtaagatttgtaaacatttgaattg  
 cggggattgtgggtgttatacatacattttaggactgcaattttttgttattttttgtattgtaaaaataacagctaat  
 aagcaggaacaagagaactaaggagggtctgtgcatttttaaacacaaatgtgaagaacttgtatataaaacaaagtaaat

FIG. 7

26/157

actataatacaaaacttccttctgaaataaaaagtagatctggtaaaaaatgtggccttttgttctgagtggtttcatttttgatt  
ttgcatgtgttttgccatattttacatttttggtcatttagaaccttaagggaaaaaaaagcgtttaccagttttcagact  
gcgtcagagggagcaacttgacagtttaacggaggtgaggggtgaaaaacccctggcaggtacctcagacagccctccaggcag  
cttgtgagccacagacacctccacacagcctttgtgtgcctacagggcagctgggcccgtgggggcagggacagatctgag  
cagtgaggtatggacagccgtctcactgctccactaaccacacccactgtctccagtggtgctgttatccctccactttgc  
gctctcctcacaggetcctgaatgtcctctaggtgtctccaaactgctaaatctcaggtgtgctgacaggtctcagccg  
tggccagtggttttctcggccttgaagttttgaggaagttactccatttggaaggccttggttgggggctcctgtgtg  
tctgcccccttttatttggctcctgcccctcagaatgggtcagagtcattgtatttcagtgacagtgacagattcccagtg  
ctagcagttccttccagaaggatgaggtcattgaaagaaggaggaaacacaagacggctcactcctgtaggggaagatggc  
agggtactagggccaggccctgaggtgaaggcaggagcctggaggaggcctgacctaccagagctcaggtctggccttaggg  
agggaaatggctggactagtggctgcaaaacagttgagaaggaccttgaaaactctccagggaagggttatgagagccacc  
aacagtaacaggatagaaacccctgtgatttgggtggtactgtggcagagacaatgctgtgtttaccaaacctcactgggccc  
acagctcatcatttccagcctcctcgcagttagggtgggaccacaaatgctgtgtttaccaaacctcactgggccc  
atgggcaccaatttccaaacgtggcacacaaatgctccacaggggactctcctacggatcacatgttaaagactggaggg  
actgtgggtccttgagtcccccttagaggccggccgctgatcagtaagagctgtgtcctggcctttgcatcactggaaag  
caaactcactctgctgatccactgagatttggggctttgttacactagttagcttttcttattctcactaatatacaaat  
cagcactttgaagtgcagtgacatgtaaccaaacccttaagtacatggcactggctcagcgtgtggccaggtggcgagc  
agtataagacaggtgcagctggcaggacaactgaggtcctgcccaggcggagcaaaacatctggtaaggccgtttcact  
cgataactcggaaaaacaaactccatgtctgctgactttttgacttgaggggaataaacaacaacaacaacaacaatatg  
tgttgattgttactggctgcttttggcgaggcatcacgagaaaaagttacttcaagaaagaattagccagtttacaagc  
agaagggaaaaagatggaaagatgttaaccagggccttgcaattctggaaaagccaactgcttctagtctccaaccagta  
agaggtgcaggtgagcaacaaaggcccccctagagacttagttgaacacagtgactgggctcatggtgacctccaatgaa  
gggtgtggtcttccactggagcatggaaactgcaaagccaccagcttgagggaggggcacgtggttaaggaaagcaaga  
aaaagctgacatgagaactgtttccagggaacaaaccatgaatgtagtactgacccctggacccctggaaataaataaa  
gatgtagtaaatttaagaaccattagcaaaagccttaaaaaaaaaaaaaaaaaaaaaaaaaaacggtggcaattgaggaa  
gtcttaaatccaccttcaggcaggaaaaatggttgtgaaagctgtgtccagcgagggtttacaagctcagatgtgggccc  
aagagcagaactgggctgagcgaggactccccccagccaggacagggccagcaggattgcaaaaacaccgcagagca  
gcaattgctgcatctgacccgctccttctgggtaggagggcctaggaggggtgctcctgctgttctcactcactacctgc  
tgggtgggaggtggattaaagctgtctgttaattcaggattgtggacttcaaagagccacctgagggagggctgcatccc  
tggaggtctgggcccgggtgtgacctcgtatgtccccctaggaggggtgctgtgacccatgtaggagaggaagcaca  
aaactggttttgggtgatcagaatagcaaaactgggcccagatgcttaccacacatttctcctcctcagtgccacagcttcgac  
tgcatggcctgggcccgtggcatctaccagatgcttgggaccagccctcaccaatgatgtgtcatttcaaggccattc  
acagtgacagtgctcctccactgtctccatcaaccttgaaggcttatggtgaagatggtaatacaacaggattgaagga  
gactgggtccctgaatgactttgtggaacagagcactataggccaaaggtagtagaattaggtatctgttttatgtgc  
ttagtgaattaatatattctctagatttattccactagtattcaacaaatttttatttgagcacatcagggctagggca  
ttgtgtcaggatcagagattttgcaatgaacaaaacagaaaaactccatgtcctccaggggagttgataataaaatgtcag  
ataagctctgtgaagaaaggtaaaagttaggggttaggggtagaaagtgtctggggggatttgctacttaacatagtggtgcca  
ggggaagccattctgctaagggtgacagctgagcagaacgatgagtgaggggggcaggaaggaatgtggaggtctgggaga  
agagcattccaaacagggatggcaggtgacagggccctgagggggcagcatgctggtaggcctgctggatattcttgag  
aatgaaaaagtccattcccagttgtttgctgggagaacctaacccagttataaacacaaaagtacttgagagctgggttaa  
tatacagccagctttccagcagtgcaactactgtgtacatcaagggaaaaactgaactcgttttctttaaacttatcat  
cagctgggtcatcattttgacaaaattctgtcaacaacagcagtgctcatcctggcatctgtatgggtcacgtctgaacaga  
cacacgccttcagccctgcaggtaccagctgtataacaagaactccctccaccctgtgtcctggaaacaagaaagcca  
ttagaccggaagatccgatggctatctcaaatgtgctggatggagtggccagggccactggcatgcccgtgaagcctt  
tcttccacgtttgggttctgccccttgaagaactccatttctgagtttgtgtgtgttttactttctagtgtgttctca  
tcttaatttttctctctctctctgcttgaactgaaggttcgcttgggtgtggagagacagggcccccagcagagcagct  
tcccagacatcctccgatccagggttcccagcagccggcgaaggcagggctgtgcttctgctcagctcacaagca  
tgccaggtcactggcaagctgtgtgtgttggggactgctcctaaagccctgcacagccctgtcctcctggccctc  
tggaaattccacccccgtgtccacatttcatgcaaaaatgagctggttctgtgagcatggcccggcctgactcgttagt  
gggcggttaagtggtttccacttcaaccttgcaacctaatcaccgggtccacaccaggatggacattcatgagccgtgaag  
tttccagtaataaatccacagatgcttccagcaacctgcttttgcacacctccactcccagccactgcccaggcaaca  
ggtaacagagacccagtcacaggagggcagtggtgggggcaggactgcagctctccaaagccctgcaacaaacgcagc  
gcccgtggcaggacaaggaggctgacattcagatgtggaggaacaaggcatgacccattcctggtcatgggggcccacagct

FIG. 7







28/157

gacctgctctaggggatctgggtctccagagaggtctcagctacaaaagtgacctctcctgacccccaccttgtgggaga  
ccccagtggtctcaaggtctgtgaataaagggaacttacaaccaagatgcctcctgacagggagagcgggggctttacagtc  
aggccccagcctacatcccttgccccaggggacggcacacaggggtgtccaaacctacaaagggcagcggagtgaagcac  
cagggccgggttggggggtaagccagcccccttctcaaggggcccagtaggcccacagaggcggccccagggcagggcaggggtgga  
gtagctggctgtgctatccgaggtcgctgtcctaatacagagcagggccgggagagccaggacaggaaagtatggagagcag  
ggagcgtctctccagggccctgctgtggaggacaccttgggggtgggagccaagtgcagacttgaggcaaaagcctgtc  
tctgccctctgctgcacccccacccccagggcccttcatgcctccacccagggcccaccagggacggaaggggtgtgaaggtg  
acagtgggggaggtgccagctcacccctccccctcctgcacctacctgcaggaggctgtgctccgggtctcctgggggtt  
cagctggtccagcaggactggggggccacccgctggcgaaggcctgaatggcaccatctatggaggggaggggtccaggggtg  
gacttagctgggaaacctagcacagggctcagcacaggggacaaagtgcagagacccccctgtgcgaagactggctcctt  
ggacagaaggtgtgtgtgagggaataacaaaataaagcagcgtgtgacagagtactcaactgtacagctctccatt  
ctccccctctaaaatggctccagccagggcggttgggagggggcagtaagcgccgccccactccccctccactcccccc  
cgggccccctcaccaagcagcggttccctggtaaaagagccccgggaaggcttcgcagtcctcacgcccgttcccgtggct  
ccgcagtcgcacccagggcgccacgcgcgcgtcacgttgtccacgtagttaggggtgacggcggtgctgcggggacccct  
gagggcgaagtcatcgccccacccggggcgagatatccctggagaacccccgcccctcgcccggtacccggcgcgcggt  
cccacagggcccggtaggcgccagggcagcgggcgccctggtccagcagggcagccgtcgggggcgctggggcgctgggg  
gcacagacactgaaaggccagggaggcgaggcctgcgcggcgggagggcggtgagccggagagaggccccgtccccacct  
cgccacggcccccgccgcccgcgcacactgcagacccggctgcgtcgccagaagtttaagggtcaaggcaggaggg  
cgggcgggggccgggccccgaaaaggcgaggggacgaaaggtctggcgccgacgctcgccgcacgcggggcccgcg  
acgggcagaaagcagtgctgggtgagcgcgggcgcccgccggcggaagaagcgcgagggccccggcgagcgggcg  
cggggacagccccctgcgcagccccggcccagggcactgcgccacatactcggagcgcaaacgctggcaccgcgcgtccgc  
cgtgcaggcttcggccgctccacacatcggttccctccgacagagctcgccgacccctgggaaaggcgcggggagggctgc  
aggctcagtcgccccgcatagcacacttaccacagcccatccacgccccggaaggtgggggacactgaaaaacccgaga  
aagcatagagtgtggcccaaaggcggggcccagctggcactgatcggtctccgtcgtggggatgggcagcgcggggtct  
ggagggctctgcaacacgtcagggatggagagggagcatggaggctgagccaggccttccccaaaagctgaagccccctcc  
ttcccaagcatcgccaccttcccagaacggccacacagctccaaagacgggtccaggggctcctcctacactcaatggc  
actacctgacgggggaacaatgaggtttctggggtatggccaatggagactttaaaccagagatggcatagaagaca  
gtggctgctcctagtctcttaatgacaggtctgtgacgccataccttctcctctacccccctccacctcgccgttatcc  
tccacctggagaccttgggtgggtctattgatttctgtttctcctccttctgacctctgctcctcctcctgccc  
cctgcatgggtgtctccctctccagcccttccctgatgctcccttgagtcgaggtcatcctgccccatctgcactggctt  
cctccagccccctccaccttcatcagccaggggagagcggtttatgcaagagctctggccactgaagaacccactgagt  
cctggctgccagggtgtggattcagagaagtccctgacgacacaccgcccgggtttgtcctgagccacaggacccaaggg  
ttcatcagtggttttagtgaccagggccccatgtaggtgacacctgtggagacagagaaggtgtctggcaccagctgc  
acacaccaggggttgggtgagacagagggtgagctctgccacatgcccaaggctgtgtctaaatagagccaggataact  
ggaagtgggacttggcagccccctcccaggatttgagtcacaaaaaccaggctaatcgaatccccccagcctccacaa  
gttctccagctctggagcatctcattcagactagaagcctttctgggtgccagcctgagtgagcgtgcacagcctggcccc  
tcatccatctcctgtgtgtttgtgacaggggctacgggggtggtatccccccccaggggtgggagcaggggtgctctct  
ggcaatggctgtgtctttcaggctgtggattggacccgctgggtgagagtgattcctgctgtgacttcagcacctgctct  
gcaggaaggatgaacatagctcgagggttagaatggtccctagtgaggggatagatggcagcggggtgagtcctccc  
ttcctccccctccccaaagggcacactttgctttctctccagccttgccaggcacatcctcaggtctcagctgcacaca  
ccagcacctgatgtgcctaagggtttgaacccaactctgctgtgtgcttacctgtcaccccttgcccagaacaccccttac  
ccagctctccacccctcacagcagctcatctcagctgcaagatctcctccaccttaaacctttaccagattccactttccac  
actacactcggtaaagtgacccctcctcccgacatccatgcagggaacggagcagtcgaagaaggtggaacaacagatg  
aacaacaccgaagattcaccaaggaatctgtgtgattgggtaggagggtggggctgaggggaggggtgagagagtgggag  
atggcagacaataggcaaaagtctcgggggacgggtgcaaggctggactcgggagggggatggcggtgggtggcttgac  
ggggctgggaggtgaggtgctcagcacatagtggggctcaattgccaggggtccccaaagcgtgcgcacaggcctctcact  
ggcaaccccttctgtggttcagcttttctgggaatatcttgccctccaccttgctcactgaggagcagggacagagaggc  
caggggttcgggggtccaggggaagagagggcgctcaccagtaacagcagcagcagcagcagcagggccccagggcagcgga  
catgctggaccttcaacagagaaggatggctggcagagtcctagtctgataggcgcccccttctagtgcctctggagca  
ggagcacagtgtaaaaccagctgtctgcgtgataccccctgggaggagcctttgctgcagcaactccccccctcattaa  
gtctccaccccggtgcacccgtcaggtaaacctggggagccacgggattcgccgctgaatgtgcctagacggaccccaaa  
cacactagccctgccacgcacacccagcaacacacagacaacacccagggagacctggtctctcccatgctaaactagg  
aacgctatgaaacagcattttttcttttcttttatataattctatttttaacttaaaaaaatttttttgaaatagagt

FIG. 7

29/157

[illegible]

FIG. 7

30/157

cactgcagtgttgaactcaggctcaagtgatcctcctgcttcagcctcctgagcagctaggcctacaggtacacaccacc  
acgccagctaatttttaaaatttttctgtagatgcgggtgtcttgctgtgttgcccaggctgggtgtggacctcctggcctc  
tggtgatcctcctaccccagcttcccaaagtgtgggattataggcatgagccaccacgcctatagccaacatgtcttct  
ttttgacttctacttttggtatcttttcttaaatggttccctctgtcccccgacacacagaatgggggagaggctgtc  
agattctgagctccagaacctcagggttagcactgggattgggggtgggggtcagggaaccacctaggggagaagacagg  
gtgggaagaaacaggaaggaagggtcccaaaattatgtttgtttgcagaggccagccaggctccaggggagtgaggactc  
agtccgaacctatagggccccaggaccactagcttctggccagcagtcacgctccacagagctgggtccgtggaaattgc  
atgtaggagacacaccagactccaggacagagcccttttgggatggccagcactaccagcctccactggtgagggagg  
tcaggggctgtgtgacctttgcttctgggactgatgggttatgtagctggagagtgtgccagcagctgtctccagcct  
caggaaacttctagtgtggctctgggttccctggagtgggtgggtcgaagctccactcggggaagaaacttccaagctgcct  
gcaggtgctggagggtccgggtgattcactggctctgcccctgcagttcaagttcctggagtggctgtcagtgggccacctgt  
ctttaaatctgttcattttaggagctacctctcaccagagcaggatcttggcatctggacttgatctgctgagaatgag  
gaggtatgttgtccctaaaggactggggccccaggctgcaagctgtgtggcagagagcccatcctcactcagtgaggac  
cagtgatccaggaaaagccacagcttctccctcccagcccagggttccagcatcctggtctccatgataaccaagag  
gtcataaactcatttccataataacctgagcccagaaacctgattagggggcagcaaactgaggggtgggagaggtggga  
gggtgggcgatgagagggggagggttgaatccagggtccctgcctaccttgggggtcaggcgagactgctggcagaggct  
tctcagggtggctgctgggtcatgagagttctcagggtctgggagaaatgggtggagggtaaatgttgtgaataggtca  
gcaggagacctggggctggggaggggcataggggactcaagggtgactgggtgctgcccacttggaaggaggcaggaggc  
atgagcccttcccttctcccttcccttccacctccccctgggtgcctcactcaccaggggcccagggtgtccagtggt  
gtggggcccaactccatggggtgaacgcgcgcaggggtgggtccctgtgtgggcatcttggggctgagcaacgtgat  
aagagtcaggagggttggcacagtgatcctgagtggttatgtcctccccgcagcatggtgtccagcccaggaggttctg  
cgtttactgagttcttggggcaccatctgctccaagtcaccctctcagctccttctgctcctctcaggggagcc  
ttgggatccagggtccaagtgagcctcatgcccctggctggcactcctctctctagtccataacttctccaggctct  
gacaccaccagcagcctggcactctccagatgctggcatcgctcagcttccaaagaaccttggatgtccgccccttcgg  
cagctatgtctgctctccttggccctgggtgcctgctgccttgatgattccaagccatctttagctgtcccatccca  
tcccccaaggccttgtcatctcctgtgatgttccctcaaaacattctccctgcccctgagactcccgctggggatgaga  
agcagccgcccacctctgcagcgcctccctcctgctgacacgcaggtctggccaccttgcctcctgcgccacagacc  
tcaatcacaactcgctttgtcagggtcctcttagctgccaccggggcccagggtgggtgcctgcctgtctgttatgcc  
ctctgcccccatctctggcccaaactcatgccatctcccttggcttgcctgggagcactcccaagaccaggctatgtcagac  
atggccacagagtgcctgcctgcctaggccctgggtgcagggtgagtcctaggacagccatgcttagtattatgtgact  
ccccactccgcccaccaccagggtcacagagaactgggttaaggcagggccctggcacaggggagccagcaccgcagctg  
accagtggtatggagtgaagatgtgctgggcccagcatttgggaacttcaaggggtgacagaggtgatttgtgcaga  
ggaagtggcaaggggccgaaaactgggtgagacagaggctggacaggcctccgggggagcagatggtagaggagctcaatc  
tgagccagggaagaaacagggtcgaagtcaagggtcgggcagcagctgggtgggagaaagcaggagagtggaacagaggagct  
gtactgggaggtagagggtccatgccttgcggtgctgggtgggtggggcagggtccacccctcgctgactgggaggccac  
tggaaacctcctgttcaagctacttcttccatggcctctggggctgctctctgcacctgggggcaagggtgagggcctg  
ccccagctccacagccccagcagagctctgaggaggggaaccgcaggagtaggctcaggaagcagggcctcggagccta  
ccactgcacgcagggtcccttctgcagccccagctgcatcgctgcagatgggtcctgggagtggtagcaacaccagg  
ccaggccggccccctgggagcagaggcagcaggacgctgaggagcatggccagcaggaagggtgcatgggtctcggggatt  
gggggaaggggcgctgagtcctgagcaggtgcaccacccccagctcctgccacatgccccactggcataacttccagcc  
tgcacaggggcaactgtccatgctgccaccaaagcctggctgtgcagaaagggtggagcccagcctggagcagagtgga  
gttatgggtgctattgcaaacctgcagagaagagaagaggagggtcacgtaggattaggaaaccccaagggtcaccctcact  
cctcgggtctctaccccggtggctgtggcaggcagtcaggcagcgctgaagctcctggaaggcatctctcctgcagcgct  
gctctggcacacctacgggcagtgaccaggcagtgagggggacactggcctgcgggattcaaacggcaaggagggggtcg  
gggtgggcagagctcaccattctagggtccacactgggtgcctggctctaccaggccagggccaagcaggtccagctgggca  
ctggggagtggcaaggctccccgacaagtcacttccctggccatctagggtgaacggtagagtccactggcaccatgtgagg  
tgagcagagggtgggttccacccctggcactgcagcttccacacagggcatccctggggaggaagttagaggggggtca  
acagctgcagtaaccccttccccaaacccactccatagcttctgctccctccactcagctccactccctacctccctgc  
acagggcaggaagtggccctcgctgtcctggccgagtttccatgagcatctccgcagagttcaccacactggaaacagg  
cctcgggagctgggtgggagcctgaggaagcatggggcaggctgggggcagctcgagaggggtcgctcagcgggct  
cagtttccctgcccattctccacagtagaaactggcccaacagaggttagggggcaggggtgcaagggtgctcgtg  
tccctcaccagggccccagagctgctggcactgctcctcagcgtgggacatgcggcatccagcagtagccactgccc  
ctggcacagggtgagccgtccagtaggtaaacgtctgggggacagtgaggagggtgcccgtgcaaaactcaggggaggtc

FIG. 7

31/157

acagt caccocatggcctggcgccacagcgctccagccggcttcagctgcgcaggtgacgggtgggtggggaaggcagagag  
aggccacgtgcagtgagaggtccatgccagagcgcggtcggagctgggggagccaggccatcccaagcccgagcacc  
acgggggaacctgaggggacccaattaaactaaggccaacaggccggctcccaagctccccgaaacctcaccctgaacct  
ccatgccctcaccaggcagcgacgcagcagctccccgtgggcgcactgggccccggggcgagcgagcagttgtgagcaa  
agcagcagaggtcgcggcactcctgggaccagaaaggcaagaaggggccaggtgagggcgagcgccccagacctgagcg  
gagagggcaagtgggggcccggcgagccgacttaacctggccaggccgcagtcacactcctcgccccctccacgaagc  
cgttcccgagagcgccccggccacgggagtcgggggtccggggcattggagaggcaagcgccgcccccttgcggaag  
aaggcgcgagctggcgggcggtgcaggcgctgaacacgcgcggaaacgggtgcctaccggcaggggagggcattggggc  
ccagagagggggaagtaaccgcgcgaaagtccacacaacaaggcggaagggaagcagtgccggccccaatagtgcagcgccg  
ggaccaaggtggaatcgcgacccgacgggtgctcctcccggtgtaggagtaacctcgccaggttactcggaataatct  
tcataccgttgagaatccactttgcctgagcttcttcccttaagcctcataaaccacctgaagcggacactatgatca  
ttatccccattttacagaagaggaaactgagggacgaccaaagaaacgcagcggaggaagtccccaggactagccgcccc  
gccgagccccgacccccccacccgcgtaccgggtggccgcagccatgacgcagcctccggactcggccgcagcctccacg  
cagcagccgtcggggtcgtggctgagggcgaggtgctggccgactcctatgggcccattgggtggctcgggcgagctggggag  
ctccgagtggtcctggggggccgtgggagggcggtcactcgccgctgagagcctcctgtctcctccgcccccgccgc  
ggggctcaccgtgctcagcctcccgagctcctcgccgcggcaccatgccctcgacggcgccaggccactgtggcgccct  
ggaaggcgcgccccctggggggcgagcgcgccgtgacaggcgggggcgggaggtgagggcgccccaccgggacccgcg  
tccgggtcagaggcaaccacgtgagcagctgcgcggagtcgtggggcggtgcgcccacagcccccgccgactgcagg  
aaggccagagcgtggcgttggcgtcctgcgtgacgcggctgcgggtcccgctcgggtccacacctccaggccggtcagcgc  
cacctgaatgtccagagtcctgagaagctgagggcgagggcggggtgaagccgggacagggcgccccctcgccgggtgg  
tccttcgtggggcgcccttcctcttccccaaacccaccagcacctgcctgtcctgcgcgcgccccaccctaccgctc  
tctcccccgccccccaaacctgggtccacgtagtggcgacttccaggagacgctgtttgggtgtgggttcaagtttcggtgc  
cgagtcagaactgggaaggcagaaatccccgtggcttgaggggctgagctggccccatccctgaaccccgccaacccctg  
gggtctcctcctcaccagggtgtgggtctgccacaatgtacagttccaggtaactccgggtcctgcgcgcttctcgccctgcc  
ctgcggaggtgcaaatggggaccctgagtggaagctgctgggcttgagccctgacccccaaacccagctcccaagagga  
gtttaacatgtttctgggaactgtttcttcagacttcaataaaaaactgggactcgaggcctgtgaattcctgtctct  
tctgatttgaggggctatagatacagcattcccactcccatccgatcgatgccctgacccctgctctggggaccaccagg  
aaggctggtcatgcccgtttgttccaggatccctgtggccacaggttcccttccagggtgagcagctgctccatccgaa  
agatctcgtgggttgagaagtccctggagccccgggggtggccagggaagcagataatagctggcattcctgctgaggggtg  
atcaggccactagggtgcagaggggtaggagcggtgtgagggagctcttccccatcccaggccagcctcctctccca  
gagctcctcatcccagagcaggtgcagaggactaccacagagtcggggaagcccttactcgcccttggtagtgga  
tgatcctagggaaggaagggggccagccccaaatctcagccagggtggagcaagaggggcaagagggaggggtgtggtaggg  
gctgggtccaacggcccccttaggaatgcaaggaggtaggggttaggaatgggtgggggggtacctctggcggtgcatccc  
agagcccatggaagcatctcacggtgtgggttgggggccagcaccactggctgcccatctgggcccgtagtggggtttctatg  
tatcctggggccagcagcctgctgagaggggggtgttacagggaacactgaattcagcttccctcctgcctcctccaggatg  
tctccagccttccctccctaaatgctaatggagcagctttatgagtgagacactcacagtgtgtcttaggggaagggaag  
gagcaatgggtgacttgctcagatcagaaactcttgggctagaggaaggagccttggtgatggcttagtggtgggagatg  
tgaatatgggaagcaccagggaagcgcggggaggtgggaatagggggaagagtttgggttccagggggacctgcag  
caggcagcaggatccacaggatcgggaggggaggtcaggagacactgccgaagaatgggacttgaggttggggaaatg  
cggtgacctccccagttccctgcctgctgcctccttgggtgggcatctgggtcgacctcttgccccacactgcctta  
gatccttgaaatatttctcagacttctagacccacatacctccacactgtccttcagtgattgatgctcaccacctg  
cctccagagaaaaagaaatcgccacctgccagctgcttccacctcctgccttctccacaccactccggtaatatgat  
tccatcttcagggtccatctcaacaggatcttccccacaggatggatcaatcataagtcaatctgtcttctttaagaa  
aatccttaacccaacctcaccttggcctcattacctccagaccacccgctaataatgatgggtgcttccccctcccaggcat  
tccaccacctgccccagctctgccccctacccctgccccacacacaccccgccacctaggaggtaggtgatgtgaccacc  
ccattgaaagggtagggaagcgtcggggaaatatggttgggcacagtggaactagagtttgggtccctgtccatccgactcca  
cgagggagaaataaaatcgtgtcaagtgtcagaacagcgctcggtcaagcactcagtaggtgatataactgataac  
ataatctgggtgggtttaagagcctgcgctccagccgggacacccacccagcccaaggcaccttccctgagcagag  
gtctgcctgtcccttccccagctcaaaaatcttggctcggtccaggacactgagcaccaaaatggccttctgtga  
tctggccctcctgggactctccaaatcattccccctgctccctccctgtagatggagaccttccaggcaggtcagac  
aaactgctgccccaaagtgtagccactgcactttttctttttctttttcttttttttttttttttttttttttttt  
ctcttttttttgagacggagctcactcttggccaggccggagtggtgcagtggtgtgatcttgggtcaccacaacctct

FIG. 7

32/157

acctccgggggttcaagcgattctcccgtctcagccccctaagtagctgggattacaggcgccgcccaccacgcctggctaa  
ttttttgtattttttagtagagacgggggtttcaccatgttggcaaggctgggtctcaaaactcctgacctcaggtgagccgcc  
cgctcggcctcccaaagccaccgcctcttgggtccctgccattcccttagcctgggtgcccgtctatcttttccctcta  
ggatttctttagactcagcatatcttgcaaatgtccactaggtgggtgctcactcatcgccagcagggagctaaacaagccg  
ctcctgggggttgggagggcgagggtgccccacagcggggctgacagcctcagcgggtcctctcagcctccagggagccaa  
ccacaggcctgcgtgactctccctgtcatctgcaccctctctgggggtcctctgcccatccagccacccgcacagatctgt  
gtcagtccttgcctcccccacactgatccctcctccagccctacccagcctggcactcactgggtctcttctccagctca  
agcaggagctcctggcctcagcctccagggccaccagcccatgtctgggtcgagacctgggcaagaaagagtggtgga  
gctgagatgggtggcctccaggcctcctgctgccaggagtaggtggcctgtggagccggctggggaggaagtctctggg  
gagaacgtgggtgggagtgagcaggacccccacatactatggaggcggtggaggagtgagaacatacaaatgtgt  
cccaaaactcaggatgttggcagctcctgacacagccacttgggaaggcggtggcacagcctgccaggccacaccagcatcc  
tccctagagaccagaggtccagaaaggtgccccctccctggccccgcctctcttctcatgccagaaggggcatcaaa  
agcagggggaagacagaggggtgctgaggacattatgggggcatcggttagccatgggtcagggcctcctcagagcctctgc  
tacctgaggtctgggttccaaatgagctgctgctcatctcctatagaattcaaatgtgactcctccacttccaattttggc  
aaactgctccctcttccaaagttttctgggctccagcagcccccgctccctccggctccgacacctgctcactggacc  
cacgaagtaaacatggacgccattccagcccaagagacacactggctcagctaggtgtcaggaggtcggctgggacgg  
ccagcctctctccttccccaccctcttggcgtctcccaacctgtgggaacacccacttcccccttgtccactcagcc  
tggctggggggccagagtgaggagccggccaggagcttctgggaggtgctgcgccttcggaatgtttaacccccgact  
ccttttctccaaaatgcactggcctggggccctgtccaagggtctcagagtcttggaggaggttcttctctcgcaagt  
ggggagcagatgggtccttgcctccctggccacaggccccacaaggcctccagcatgagctcatgaggtcggatgccact  
tgctttattggggaaaggtctgcacgggaaaaaaggccatactcgaggtccctgttctctgcagccctgctatcttta  
ctcttgccctcctggtaacctgcccccttgatatataccccctcatcttgaaatgtgagtggttctcctgcttctggaggga  
tacctagcctctactcttctctgtaccatcttggcaggcttctgggggcagggggccaccgggtgggggaagcagagc  
ccctttggggctctcctcttgggtcacagcccaggccagacagacaggagggccagaggcagagtgacccagtggtgt  
ccagccttccccctcctgggtatggggagggcaatctcaagctcaggccagtgccgtgcttgaccagtgggaatgggggccc  
ttatgggctcaggggatccagtgagggccctgggttgggagctgctgggtctctgggggctctcagccttcatggcaa  
tgctccccctgcttccctcttgcctggatttggacagtaggggtgaaaattccaaacaaagagggtcctcaggaggggca  
gggggtgtagccaatgggtttaaatacgttcagaccttagtgggtcctcaggctccagcctaaagagctgtgtgacctgga  
caatttccccaaagctctctgggttccgtttgccccctctgtaaaatgagcatatcaaggctactgccccctctagtgtgca  
gcacagatattatggcacaaacagatggggcatgggtattctggaagcgtgtgaagagcgggattgggaagaggctgggg  
cagagcgtcctgcagaagaagacacatgggggtggtcttacatctgggggacatcaggagagtgaccactgcccccccata  
ccagaagtggattccacaggagccagtgaggtgaaggtcaggccttctgtggcagggccctgagagggacagcagtggt  
tccacaggttcacatgttctgtcaactttgcaaaaaggtttcttcttgggtgcttttttttttttttttttttagagg  
ctcctgaaaagcttcaggacccacaaactctggaccatttctgcctgggtgggggtgggggtggccagatcatccaggg  
agggaggggaaagagggaggtgggggtggagaaagctgaaatgacttccatgtgtgcgggctcacgagatccagatgtccaa  
acccagtgcccttcttctgcccacttgaggggaggggagggaggggctataggagtagtgacttgggtggtctggggga  
ccccagcaaaactagaagctgtaattgtaggagagagacaaaagggtgggaggttcaggggccctgtggagggcggggaga  
catggcactgaccggctcctccaggtgacgggtgcgccagggttgctccatccaggacccagtgccgggtgactggctgcc  
caggggatatgtcctggagttaaagacagagcacagggtgaggggacctgaggaacacaggggcatgggacaaagcagagg  
gaggggggttagaggacatccccaggagggcactggaggccttttggggcagacttcaccttcaacacgcgtgggtcagc  
ctggagaaggagggagcggcctgggcatccttggatctgaggagctatcaaggaggaggaagagaaggctggaaggg  
acagctcagctggggacacgggagtccttgaccttgtcgggggcaggcttgggtcggcgatcacaaggaagaggcc  
aaggccgcagtgacagaggggaaggaaagagcgcggcagccttagggatttttagatgggcagcagatgccccttaggggtg  
agagatgtacgaagagaggacacttgtgcccccccatcatctagaaaaaacaacagccagatgtgcttgccttgcaggtcc  
accttggccagagctccctcggggactctgtcctggggcagggttttgggtaccttggccagaaggccctcctcatct  
cttcaaggggaggggaagccttccggacggagccttgggtgctcctggccgggtgtgcctaaaggggctctaggaggaatc  
ccagagccaagcattactcagaggggcctgggaatgttccctggaatgctccagccctccactggccccaaccactc  
tcacaggcccgccctgcaggagccaggcccccaggcaaccagagcctgcagcagcctccttccccgggtaccagtcacca  
gctccagaaacagacagcctccccctccacgcagccctggcctcagtcctgctgggtgatgggtgctgtggaagtga  
ctcagctcctgctaggccaccccaactccttttctcctccacttctcctccagactacaaacatcaaagaccttcc  
tccaagaagccctccttgattggatgagtgaatttgcctcaggcagatgagggccgagaggaggtctgcacacttggaaag  
gaggttagaggggagtgagggaggggtctgagtggtgtgggggaggggaaggaggggaggtctctcagccagaga  
gcacttaactgagagtagagaaccaagcttctgctgctcctaggcctctaagggtttggggaaagaggtaggggtgggcccgg

FIG. 7

33/157

gcacaggtgtggtgtgggtgcagtggtgtgtgggtgtgtccacatggccttgccgcgcagtggtggccacggggcacc  
ctgaccccaatgagggagagagggggcagagctggagctggagctggagctccggtgaccgggtgaatgggggtggaaccc  
gagggagccaggctggtattgggcacatagacgcccctctcccaggggtcccatcacctcccctgaccccaggatagggc  
tcagaggggagggagcagtgagccgctggggccctcccctggggccagaacagaccaggccctgtacctgttgggtcc  
ccacacagtgctgtggaagccaccgcccagtgctgcatagcacagcccagcccgcagtcgccctcccctgggtgcccctccc  
tgttcccggccaggcacttgctgtgcaggactggctaactcccaccccgttgccagaggttgttccagcccatcttaa  
catctttgtttggaggggttaccgagggagagacagtgctgttccagagcactgctaaacagacacctctctatctgg  
agaggcccttctctatctcaccaacaaggcaacaataataaacacatacactgcccctgctgcccctgggaggagggac  
gaggggtgagcaggggtggaggccacagctagtctgtcagcctgagagcaaagcaggagactctgggggactcttgggcatg  
ggggcttccctagaggatggagcccgcctgagtgcttaaggggtggaggagcaggagcgggtcacacggtggcctgcccgtg  
gaagctggttgtgagagcgagaatccaggcagagggggctacggctatgggctgggggctgggggctgggctgtcccga  
gggggagggagccatgcccctctgctttgcccagcgagtgccagccgggcagtggtgggcaagtccgggcccggggccccc  
caagcacacttgagcctcccggcaggtcccagggagaccccccaagagtgcccacgcccctgacctactggccgtat  
gggtgcccggggcgtgagacctccgcccgtgacccgagctctgagcagaacccatcccggccaccaccacggcgcctag  
cctgcccctcagggcgaccccgcgcgcgtcctcacctgaagcaccggcgccctggcactggccagagcagcagcagc  
agtagcagcagcagcaacgggggtccccgagctctccggggcctccagcccatagctgtgagctcctcgccctctaggca  
ggggctcgcaactccggctccgcccaggctggattgcccgcgaccgctgcccgggtgcagcctcaggccgcccgcctccgga  
ccttcccggcccccacctcccacccgcccgcctcctcgctcccgcctcccctcccggccaaaccccgcctcggagcctggccagg  
gccccgacggcgccgcccattggggggagccgggtcgccactcccggaccgcccgcctcgaggggggtggagctgggcggag  
gaggggaatccgtgcccgcctcggatgacccggcccgagccgtcctcccgtogggtctcagagggcctctactcctgaga  
ggaggagagaacccgctgggaaggttcttggaggacccgcccgtggtgggatgagggcgtgggcaaggccgcctctcgct  
gctgaagttggccccaggagcgcgctctccgtgggtctcctggggccgactctctgtcccctccttgctacccgtcctgcc  
ccgaggggtgcccctggcggaggttgagtcgggtcatccactgcaactgggtgcccccaaggataggaaggttcaggcaacc  
ggctcccgcctgtcttgggggcttcatgtgctgggcaaggcgatgcagcagacggagacaaccttcttcccgtggcggtgg  
ccagaggggcagaattgcataaaagctgcagactcccaggcctgggagacccttccggcctcagtaacatctgtttcatgt  
tttaaaacttttgttttctactcggtgcaaatttggatgagatgttaacttttttttttttttttttttttgagatggagt  
ctcccctctgtcccaggctggagtgcagcggcgcgatcttgggtcactgcaacctccgactcccgtggtcaagcgattct  
cctgcccagcctcccagtagctgggactacaggccgcccgtaccaccccagctaacttttgtatttttagcagagac  
gaggtttccacttttggccaggatgggtctcaactcctgctgtgatccaccccgcctcggcctctcaaaagcgtggg  
attacaggcatgagccaccgcccgcggagatgttaacttttaagcaaatcttttttttttttttttttttttgagac  
agagtttctctctgtttaccagactggagtgcattggcatgatctgggctcactgcaacctctgcccctccagattcaag  
tgattcttctgcccagcctcccagtagctggcattacaggcattccgccaccacgcctggctaattttgtatttttagt  
agagatgggggttttcccatgttgggtcaggctgggtctgaactcccgaacctcaggtagctgcccgcctcggcctcccaa  
gcgctggaattacaggcgtgagacaccgaccagcctacttttaagtaaatctatttgttttttgagaatttggaaatga  
gtaatttgggttagtgaaagttcgagcagtgagagaaacctacattcacatatctcaaaatcaaaaagtagagaaagcata  
gggaaaagtctccgtgctcttagccctcctcaccaacaggaaaccaatattgattagtttcttctcataggcttttagatta  
ttttttcacactcaagacaatacagacatatttttttctcttattaaacttttctgacttttgattttctttttttttt  
tggtcgcttaatacaccttagatatcagtgcttttagaggggtccttggtgttcttatgattattatttagagacagggtc  
tcactctgtcacccacgctagaggacagtgccctgatcatgcccattgagccttgaaatcctgggctcaaggatcct  
cccactcagcctcctgagtagctggaactadaggcacacggcaccaggccagctaaaatttttaatttttctgtagac  
agggggtctcactttgtttcccaggctggtctcaaaactcctgggtctggccaggcgagtgctcatgcccgtaatccca  
gcactttgggaggccgaggcgggagatcactggaggtcaggagtcaagaccagctctggccaacatggtgaaaccccat  
ctctactaaaaataaaaaattagccgggcatgggtgggtgagcgcctgtagttccagctacttgggaggctgaggcaggaa  
aatcgcttgaaactcagaaggtggaggttgagcagagccgagatcatgcccattgcaactccagcctgggcaacaagagcgaa  
actccgtctcaaaaaataaaaaataaaaaaagaactcctgatcttaagtgtatcctcctgcccagctctcctcaaatcg  
ctggaattacaggagtgagtcaccacagctgtccagctacgagattattactattattactactttggatttttcaaatc  
aacttcatttaaggtataaatttacacacaataaaaatgcacttttttaagtggccagtaagatgagtttgcataagtgat  
ataactacataagcatcactataatgcagacacattccctcactcacagaaagagccctgtgcccctccagccaaacttc  
cccactcccaccccagacagccactgatctgttgttctctgtctatagataagttttgcccgttctagaatttctatata  
aatggaatcatggcgatgcactctctgtgtctggttccctcccctcttccgatgtttttgagattcatttacactat  
tttgcatatcaatagtttgttccctcgtattgctgaatagtggtcggtgggttgaggggaaccacagtttctctactcacc  
agtgccacatagggttattttccagttaggggctcttataattggaactatatttgacagagagagagaggaagaaa  
gaggagagagatatttattatagcaattgggtcacgtgattatggaggccaaaaggtcccgaatctgccatctgcaag

FIG. 7

34/157

ctggagaacgaggaaagccagtggtgtgattcagtttgagttcaaaggcctgagaaccaggagcaccagtatggaggtgg  
ctcgagctcagaacaagtggggacaggaaagcagagcagcaccagagcagccctcagcgacacctcttcagtaaag  
caaggctgaacacagagggggtggcttcagtggtgtcaggtacagaaggcagctcgaggagctactctggcgttctt  
gcttactggtattcttacctcgaactggccaactcctacttaactgcaggccatggctttaatgtcctgtcatcagag  
gctgtcccttacccaaagccaggttagcatccctgactgacacttctcctgcaacaggtttcagaaggccctgtagtc  
gtccacttccctgtctctctcccaagctcctgagctccatgtggtctgggaatatgtgtgttgcctcacttccagcaca  
gtcagtgctaataactgactgtagaggggacacagctcgaaggccacatggggatcagagtcaccttacacagttgaca  
cctcccaaacccagatgagctgtgtccaagtgcaggtcagaggaattttctgcgaagtctctgagaaaggggttattta  
cattttgaggttgcaggggaggagatgaggccatcaaaccaaagctgaggaagagggtatcctaggatgcaccgagcagct  
ccggggggtcctgacagcacctgggaaagatggcttctccactggcttgttggcgtcaccctccagaggggcatcaggaa  
atgtcctgggaaccaggcaaacagtgagcattaaccttagaagtgttggcatgggtgacacccaccatctgtaaca  
cgacttctcccaaggagtgcgcagaaacaggatgtctgagggaggcactccgactccagccttcagagatcgccagggtg  
gcacttgggtgacgacaggctgatgttgggtgccccagaaaaggtctgtgtgtgaatggggggcccaagccaacgctt  
catcctgacagccctgggtgcatttagaggggaacttttgtcccttggcaagggtgggtggaatttcagggttcatagggca  
agggtatttttagctttaatagatatgttcaaacagttttccaaagtcattgtacacactctgtgattctacttatgtaa  
gtttaaaaacaggcaaatcaaatctatgggtgttcgaagtcaagacagtagttacccctgtggggggtgcaactggtag  
agtgtaaagggggactgtaggatgggtctatttcttgatctgggtgtgttgccttttggaaaagtcccttgagttgcattta  
taatgtgtgaactttctgtatgttacactttaattgaatgtacaaaaagtcctcaggaggcctcagaccactggaagcgg  
acacaactaaacccctctgagagcctccaatccaagatggacatatgtcccttgggaagtatgcagaagcaggtgaagact  
cctaagccggatattcccaaatccccccagtagccgcagcttcagcagctgcttatgggtcctccctacacccctctctcc  
ccagacagcccccaaacatctgggtgcatttgacttgcctctctccctgtcccacctctggatttagtccatgttctccac  
cctccccactgtcagcaatgtagacaagacaaacgcttagttcacgtgcccacctactgcgtgccatgcacgggggtgg  
cattgtgggtggcaaatgtgagcaacacacgaagcctcaaggagcagaaagggacacaaatcacttcagcgttaaggta  
ttgtgataaatgtcatgttaacttgcagcccttggccctctacagatgggtgtctaagaataaacccccactaacatgtg  
actcctctgttctagcccagctgtttgggttgcagaaagagactcactccagttgcgtcataggatggagttttattgg  
gaggacattctggacggggtcccagcaagagtctggcaatggagcatgaaaatgaatgagcctggaatgagggagggtgc  
ggcctcgggtacacaaagtccaccggcccccaactgctccagcgtgttagctcctgtggcaactccccacttctctc  
tctgtttttcaaccctaatcctagagcagaggggtcctcgctcctcacaccatctccaccagggtgcagggggagtcact  
agccactcaagagctctctcctgttgggtccctgatcgcaggggcaogtcatccttcttggtaactcatctctcttaa  
cagcccgacacaggggtctccaggaaaaggacaggagctcacacagtcattgaacagagatgcattcggagcaggaataagg  
atgctgacaacatctgtgtctgcccctccactcttgaagtgcagtaaggtagtaccagtcctcgggggtagggggggt  
gggtgggtgggaatgagggcatctctcttccaggggcccaactctggcaagagatgtctgcttaggcaactccactgctg  
gcatgctctctccctaggaaaacaccaaactttttatttccctcagtccttagtgtgagagtgatcacctcttccaggaa  
gccccacaaccagaatggccaggggacacccctggccctcagtagatgggcaactgagacaaaacatggcctgctggtggca  
ttcccaacaatgtcaaaagtctcagggatgagctcacagttgggagtccttgggaagtccaattccaaatgtcctctgga  
ctcaaaataaaaaactacatttccagtcctccctgtagcttgcagtagccatgtgacaatactccagccaatgggaagt  
agtggaaaagtccctgtgcagcttctgagttgcgtccaccatgggcatgggtgtgcttctacctccactcttccctctct  
ggctggagggtcagtggtgacaaagcggggccatgggtgagcctcacaatcaaaggcatcattttagggtatagagaaaaggaa  
gatagaaggatcctgagcccccaacatcataaacctcgtaacagccagactagcgtgggagagaaaaataaaattctatc  
atgttgaagtcaactgtatttgggtctctgttagagcaggctgactgatactctgttgaatacagaatgccttggtagcc  
tctgaggatgcaggatgctgcaatccaaccaaagtccagggtgactttcgaggagaaggatgtgaaagggcagggtctcc  
tttgaggggagaaatggagagaagggaggagatccagagaagagagagcagatcagctcagtgctcttcagtgctcctcca  
gggtcaacacccctcctgggcagacctctcaccacccatttttgggggggcaagaggcttggggccagggtcataaaaagt  
agatgtcacagagctgttttctgtagaggcgggcaggactcaacactgcctcattcacattcoataggctggagtcacaca  
gagtcctggccactttctcctcggaggggcagggaacacccactgggtcagcccagggtggggcagaggttgaggtctcaca  
tgtggctgcacaggatcaatgagcttcccacagagaaaacccctgggtcagaggagcctctgggagccagggcaggccc  
caggccccagcctcaggtgggttcccaagaggatgggtgggagtgatgaggtgggtgagcttggagccctggcc  
taccaggtgaatcagctgcagaagggaacacacacagggggaggagaaactgactttcacttctgcccaggtgacgggcag  
gcttgggagtgagggaagaggcctcagaacaaagggtgggggagtcagaacatacgtgtcttggaaatggaggggatgctgc  
caagtcagaaagtctccctagtctctaaagccagccgagctgacacagggattccctctgccccgagcagaacagg  
gctctcctatctcctgaggggtcctgggtcctgcacactctcccaacccctccttgggtcccaggacccatcctctaa  
atgacagcagtcctccagatgccagttggcactgtgggttagctgggtattatctgcccgtgtgtgtgaatggatacc  
tgctgtgtctcttctgaaaagccatctccaccgaattctcggccatgctctgcttacaacacagccctcctctgcaagg

FIG. 7



35/157

cccgagtggaactcagagagcctctccctccccacctcctgccccattcctctcccgctccccacaggctggctccag  
accacaagagcgtgggaacctcatggagagtgggtgcttctctgtcttctcctccctagcccttgcttttagcacagccagg  
gcagagaggaggaaagaaggaaactacagcaaggaggatttagggacaattctagaagggtttccaatttggaggggca  
gtctgggaagtaggcagcctgtgaacttaataaggaggtcctaaaggacctgggggaatttggggagaagacgaggtgggc  
tgcttaccccaatcttctggcccaactcccagctccagacccacacctccaggtgtagcaggccccaggcccaacagcagg  
agcaggaggccccaccagcagtcacaggacccagagcagctgctggaactgatgcaggcgggtgcagggtggaacacagag  
cgggactcagagcagccacagctgcaggccccagtgcttcttccaaggggacccacccaagatgacaccttctaactt  
ttgctttatcgctgaagctctccacacaagggtgccccagcaatcagtttagagtgcacaggcaaatcatgctgcagca  
ggagggtacaggggaagggaagcctttgggcttaagagctgggcttttgtgcagagtggcctgggagatggagccccctcc  
cctcacctctgacccccaaagttaggtggaggttagaggcagagcccaggacatttgaggcagaacagatgtattccccctgg  
tcgctgggctcagcgcctcgatgtccacctcagggcattgggggaagccaggacatggatgtgtggctctgcaggagc  
acctggggctgactggaggccaccagtgcactgccaaggtggagagtcagggtggcgagcggctcgctgtccactcggc  
aatccaggatgccccggaggccacctcaggctccacgaagaccatcatgggtggcgctctgggagggtctgtggggagg  
aagggtgtgtgtgattgcagaaaatgaccacaaatttctccctccctgtacctatgcacagtgactttgctgtcctcc  
cattaagaggcagaacctatctcctcactccttgaatctgtgactcagtttgaccaattgaagaaagaagtgtgtgtg  
caacttcagagccagacctcaagaggccttgtagcttctgtcttacttgttggaacacaacaatgtgggaagcccaggcc  
agcctgtgggcacatgtggcccaactggcagccaggagcgtgcagtcactctgagaccactgggcaactgcagccacatg  
agggtgtaacaaacagaagaactcccagctgagccccacagcaattgagcagataaacatgacttttgtttgaagcag  
taagttttgttgtgtcaatagcaatagctgactgatacagtgtaggccttggtgaaggctggagtgggagaccaagact  
gatgagggttatcggtgcccagacacactcagcagcgtttatctcttttgggtttttctctctttctttctttttt  
tttttttttttttgagacagagtctcagctctgtcaccaggtggagtgctgtgtgtgatcttggtcactgcaacct  
ccacctcccaggctcaagtgagctcctgctcagcctcctgaatagctgggattacaggtgtgtgcttcatgcccgtgc  
taattcttgtatttttagtagagacggagtttcgccaatttgaccaggctagctctgaactcctgacctcaagtgtccg  
cccacctcagcctctcaagtgctgggattacaggcacaagccctgcaccagccatttatcttatacaacttgccagg  
atgaggcgaatgatgtattccccattttacaaggctcacccaaaccagacagcgggagagccaggattccaaccaggactc  
aatatccgtgaagccttcaactcttggcagacagaagaggaggaaagaaactgaggtgggctgtgagatgtggggagactaa  
ggcggaggagggggttcacactcacagagcacacggagcagtgactggagcagaggcagcagccccagtggggagctcagc  
caggcagtggtacatcccagcttgagcacgagccacgtgggtgaaggcgagagtgggcacaggctccgctgcagccgccc  
ggctattccagaacctgcaaagggtggagttgcccacaggccacgggccaaccaggaggcggcagctgaggttcaggggca  
ggccctcaggcacgtccaggccaggctctgcccaccacgcgtgcacctgcccggcggaggatagagagatgattggggatc  
tgtaggccttgggggctggagcctctgggtgggacctctgaggggctctgcacattgtgggaaggctctcagctcgaact  
gggtatagggttttccaagggtggaaacctgccccctccagtgggagctgtgccccctccaccagattaggcttctcaagge  
agagcttctcagaccagacagcccgtctagtgggagctccagccctctcgtggacttggacctgagagcatcagagc  
ccctcctcctccctctcttccactcaactctacctgcaacctgagtggtgctgattgagccagcaagttttggggc  
tgtgcaaacatagggtgtcatcactgcaggcacatctgcacctgcagccgtaggggcttccgggccacctggacatggc  
ctgtccctgatgccaaagctgtggaccccgtctcgtggccagcaccttgccatcatgagataggggccagctcagcaggt  
ggctcactgtccacagtgcaactgtatcacagccatggatctggccctggagtcocggaaggaggacaggacagcgtcctg  
aggggcatctgtgggcaggcagggcacagatggggaccttgcttaggcacctgtactgctcattgacctagctgcctggg  
gttggctcaggctgaggcctctggaccaatggccacttctagaagtgcaccttcccaggagtcctgtgggcagagtgtc  
ccaggcagggttataggactgtctcttctctagcacctatgtcctcttccccaaactgccatctcctgggcccactg  
agtcttggcccactcccaccagagcccagaagcaccctccaccacctcctcctgggcagccctggccaaggacctct  
gctcacagaggacttgaggggcagcaggacgggagctgcccgtgcccgggcatcctgggctggcaagagtaggcgct  
gcatgagcccgctggccaccagggaatgagagtggagcagctggacctcctgcagccaaogacctgtgtgttaccagt  
atagagtgtgggtgctgggagcagggtccgcacaggctcactgtgatgggggaccttcaggcacggcagcctccggag  
ccaggatcacccgcacacctgtgccaaggaggccaggatcagccgggcccagccggacacacagtgggcttccatccca  
gtgggttggcctaggccctgcttacctccagccgagctccagggaacgtgttggcctggcccagagggtgcccgggca  
gagcagctgtagaaacctcatcctgggctgtggccctgcagctccaggcgagggtgttggggacagaggctgctgt  
cgaggaggccaagaggcgaccggcgtggctgagggccagctgggcccggcgggctgtccacagtgccacagtaccagg  
ccagctgcccgcctatggctctccaggaggtaggctcaggcgaggttgccgggcccgtctgcaggggcatgagaggctaca  
cggagtccaggcggcagcctgcaggaggccagtttgaggtggcattcaaactcaggagggcccctggctccccacct  
atggctcccaaccaaccagctgtaggcactgtcctctgcccagacaggactcaccacagcggccctactcttaagtct  
ccttcccagctgctccactagctagggtaccaggttccccactggacacctgtcgggttccggccatgcccgtgatccc  
tgtgggcttctgtcattcctgtgagtcacacctgcatccagccagactcacagaggagctccaagggtgataggtctgga

FIG. 7



36/157

gaggcggggtgcccgaccagggggggccacacgcgagcggtaggaggtggcatccctgactgtgacgttgggcagggggga  
tggagtgggcatccaggcgctgctgcccacctcctgggtaccatgtgtagggtgagctgggcccgggtgagtggtccacacaagg  
caggtcaggttcaccagctgcccctcccgacggtagccccggccacacotgcacattcacagctggggagagggaggg  
cacaggacaccagtgaggggtcttgaggctgtgtacagcaggccacotgtctccatgtgggtgagctactcctactgctgg  
ggagccccttggccttgggagccttgggtggccacacctataaatgtggcttagaactcaagcttgggaatagcccactg  
gctcctgagtggtcccaagtagagttccacagagtcctggaaagccctgagccctccctgacctggctgggggcaaat  
ggggagaccacagacctcccctcccctgctgcactttctcctgaaattccctgttttagttttatccactgggcatacttg  
taaaatttgggtgtttataaaagggttctcctaagaaactcaagcttgaggtccacccatcgttaggaagggtgggaatgggg  
actattcccgtgccccaggcttctagaacctgtcccacctcctctcctcttttaaaggacacacacacacacacacac  
acacacacacacacacacacacactgaccttgagcgtcgaagttagctgagggcaggccctggcccagggtgttggaggc  
ctcacagatatagacacctcatcctccagcatagccccgtggatctccagacgcagtggttgggggcccacagccacat  
gcagcctgggagagctgcttccgggtccccccacaccttgtaggggtggaggccacaaggcgatcccgtggagcagccgc  
agctgggcccggagggtcactgtccacacggccacaggaggaggccagctcgtccaggacctgtgtccatcagggtagttag  
tgtgacgtggcgtggggcatctacacagggtggggagtggtcaggaccagccagggggggtccctcatccagggtcctgt  
catgtgacacagcccaggactgggggactgcattccttccccacacctggatgggacaaaagtccttacaggacacgtg  
gaggctgatgggtgcagctaggctcgtgggtggctgagcctggggcctgggcttggcaatgataggccccagcttgtgtca  
aagttatgggtgcaaagcggagcgtggccgaggtcgactcctggagggtggtccatcccgataccaacgataggggtc  
ccctctgggaactcctgtgggtacctggcagctcaggaccaagcctggccctcttgagctcaggtgatggtgacacctg  
gacccaggctcctgcagggggaaaaccaagagcaggtgagggctctccaccacacctctcacagctctgggacctgtgctg  
gtccacctaggactacacaacccaacccccatgtgctggagcgcagcccttccagaccaccgcccaggggcccactcagcc  
ctatgccttggccccctcctgctccccactctcctgcacaggctccatcctggcatttacctcccaccacctgggtgggatg  
gttttactgtgtcagtcatactcgcctaccaagcgttagccttatgggcactggacttgtgtgacctgtgtaacccaggt  
tccaggcactgtgctgcatgtgttagctcctgacaaatgttgggtccaaccataaaagtaactgaaaaggaggaggttag  
atcatcccaggggccattgttccctaactaggatctggactcagcatgggtgaccgagggttggaaagggaatgttagcagct  
atgttatcttgagcgtctcgcaggccagccctgcacccagactccaggccacaagagcacaggacctggggaccaggcc  
aaagggcagctcctcagtgggcccctgggtgtgaagcaggagaaccccagggttgccaggagtgagatggaggggacagctgg  
aatgctaaagcaagaacacagacctgacctgagaccacagcttagcagggtcacgcaactggggaaaagctcctaaaactca  
aaccttcaatttcttccctgcgaaatagggtatgctaagtaagaaaagatacacagaaccaaggctcaaagtaaatgtccca  
aaattgggtgctgtcgggtcaccgtgggtggaagaaacactcatgttcccttgaccttcgcattatgggcattatcttccct  
tttccctgggttatttttagggagaatttttttttttaagacagagctcactctgtcatccaggcttagagtgaagta  
caattttgggtcactgcaacctccatctcccagggttcaaagcattcgccctgaccttagcctcccagtaactgggattaca  
aatgcccaccaccacacctagctaattttttagtatttttagtagagtggtggccagggtgggtcttgaactcctgacctcaa  
gtgatctgcccgcctcagcctcccaaagtgtggtgattacaggcatgagccatggtgcccggcctgaggggaggttttta  
tagagactatataatatatac  
agaatatgaatatattatgcataagaatacacataatgtatatattttagtatgtgtatatactcatacataaacatatatgca  
taatataaatatacatatgtatgagaatgtttactatgtaacaaaattatatgatcgcaaaacaagttatttcagaaagt  
gctcccaccaagcaagccactttgtgaattgccattgcctcctagtggctgctgcagttattacagggtgaaaatatctag  
ctggaggcaaaggaggcccttctgctgggtgggtggaaactgccttaccaggactgaaaaagacatatcttgcacaaa  
attcacagcaccagcccaaagtctgaccagagtagtccccagggtcagagcttaccatattctcagaagactgacac  
tcacctoggacctggaagaagagtgaggtgttggatgatccaagaacatttgtggcctcacagcggtagctgccagagtc  
cccaaggcccagttctcggaaccttaacttcaggggagtgccctcagccttagcctggaaaccgacctgggatgggacct  
ggggacccagggtggtggccaggaggtgctccccatggaacaaggccagcaaggccaggggggcggtgtccacagtgcag  
atgaacagagccatgtggccctggcccatgtctaggagggtgacagcttggacgggtccgggggatctgcaggaaacaga  
gggagctgaggccaccagcctagctcaccacattacaactgccacactatgtccccacctcctgccacacacacagcc  
taagccatgctcttctccccaaacccagcccgccctgcttctccccagaccacctccacctccaagtccca  
ggctgcccattgacagtcacccacagaggacactcaggagcagggagtgagagctggggcaccagcctcagtcaggatgc  
ggcaggcgtaaaaggcagcatcagttctggccacgggcagcagtgacaggtctccaggggacccctggggccacagcacc  
ccatttcggaaccaggagaagtttagcagggtgcccagcagcttcccggctcacgttgcaagtcaagttggcttctgtgcc  
ctcctgaagtgtgtgtgatggtgcaatggccaggacagtggtggagagcaggcggccacagcttactgaccacccccggc  
ccccaggagccagggtccagccgcccagcctggaaggagcaggaaggaggccccctggggtgcccacagggttggggag  
aaaagcaagctagctcacagggcagcctaggagaagtgggcccctggggactgtggggccctgggacagaggggtgcag  
tacagggaaggaggaacaggcagcccaactgtcccagctggggcctcctcctcagagaccagctgtgttgcctatcccc  
ggcctgaacagagatcatgggaaggagcagctccgctctcctgaaatgtggaggagggcagcgaaggggccccactgctg

FIG. 7

37/157

cacagagtgggttttgcctgattagatcctcctcggagcagaacaggtccagagctccagccctgccccaggccacccat  
tccttgctgactcaccctggccattgaaggtggctgaggtggagcggtgcccagggcatttgctggcctcacagaggta  
caagccctcctctccagcaagggttgtgaatctccacacgcagcaagttgggggctttgggtgaccttcatgctggg  
aacagccccacaggtgctgcagccacccctgatggcagggaagtggccacaacacggctccttgaggagcagctgcagc  
ctggcgggggggtcgctgtccacacgggcacaaaaggagccctcgccgtccagcccgccccagcggcatcaaggctcag  
cctgggtggtgaatgttgggtgtcgaggggggtctgcaggagggaagaacatgggcactcatccacggatgctccagggc  
cccacaagcctggctaggctcccagaatacaactggacaaaggcagatccgaaggctgccccaaagcagctgtgcagactg  
tgactgcacaagggcagccagaccagggtgggtgggaactcaagctaggctcatgctccccaagctaagctgtgcctga  
tgctactgagctctccagagaaggagcaccatcttctagctcagacaaagggtgctgtatgggctggctgccacctg  
cataagtcccatctgctgctgagcagggtgccacctagcaatcctgggggtcatgctgtctgctccacctgcacc  
aatgccttaccacggagcctcccctcaagatctgcccctgcagcttatctgtgatctccaagccagctcctgtgatactag  
cgaatgcatagaatggcctcttctgaaggctctctggagacagttggccaaaagagagcaatcagccaaggccccatgt  
gatctcactcagacaccagaaaactttgcagaggcagctgccctctgacagttccgggacgtggctccggcacacatgtg  
tcaggggaagcctctggaatcaggcccttggtcttacacggagcctcccgcctggcggtgagaggggctgtctgtcc  
ctgtgctgaggggctttgccttggtgtctatgggagcccagcacagcctgtggggcatgtgcacaaatagaggagtggtg  
tgtcccagggtcaccatgaggacaaagcatggatgtggtgccatgctgagcttgaggctgggacagaggagccaca  
gggctggcattggagggtaaagacatggggaggttagagaaggccccagcctcatcttctaaacttcagccccagtgga  
agaccactgtggcagggtctgtgacccagccctccgatataagccttggggcagggcaggacagggtggggggccta  
gaggaggtggggtggctgggtgggtccggcaggaggctgcaacagggtgggtggctgctcacagagcacagtgagaacagct  
ggcgaaggaggggccactggcactgtggcctcccgggcccggcagtggtatgagccggcgtcagtgctggaggccggcg  
gagcaggaggctgctgcccggaccctcgtgaagcagggtccattcagggtaccaggagaagcgggcatacagggtgtggggc  
ttaggccgtctctgcagctcagtgctcactgctccttccaccactcggtgcccggctgatggaggagcggcggt  
gcggggagagggaaggctgggaaggctcctcctcacaacccacatgctgccctatagaaagcctccagggttct  
ccttccccacactactgtaggtagctctgggctttgtggtgcatcaggagggtttgcccttaatgcaaatcagatct  
attggtagtagatggctgcagcatgggtgaagattcagagccagagcccaggcttatgtccaacacctggcttggctggg  
catggtagctcatgcctgtaattcttagcactttgggagactgaggccagaaggatcgtttgaggccaggagttccagacca  
gcctaggcaacatagtgaagctccatctcaacaatttttttttttttgagactgagctcactctgtctccaggctgg  
agtgagctgggtgtagctcctggctcactgcaacctcgcctcccagggtcaagcgattctcctgctcagctcccaagta  
gctgggattacagggtgtgccaccaacccagctaattgtatacatgtagtagagatagggtttcaccatgttagccaggc  
agatctcgaactccgacctctgggtgatccaccgcctcagcctccaaagtgtgggattacagggtgtgagccactgtg  
gccagcttt  
atctgagctcgctacaacctctgcctcccagggtcaagcaattatcctgctcagcctccctagtagctgggaccacagg  
tgtgcccaccacaccgggctaattttttgtatttttttagtgagatgggggtttcaccacgttgccagggtgctcctcagg  
tccagctcaggtgatctggcctcctcggcctcccaagtgtggaattacaggcatgagccaccatgctggccacaa  
tttttttttttttaattagctgggtgtggtgacatgggcctgagctcagctattttgggagggtgagatgggaggatggct  
tgagcccaggagtttgaggctgcagtgagccatgaacataccatgcactccggcctgggcaacagagcaacacccctatc  
tcaaaaaacaaaaagaaaaacctggcttgatcaattagctaccatgcccctcaggaggagggaaggacagtgcacataccg  
aaagtgtgaagaccgtacttttctttttctttttcttt  
aggctggagtgagtgagtgctatcttggccactacaactccacctcctgggttcaagcgattctcctgcttagcctc  
ccaggtagctgggactacagggaactcaccacacctcgcctcccagggttcaagcgattctcctgcttagcctc  
ttggccaggatggtctcgatctcttgacctcgtgatccaccacctcggcctcccaagtgtggtgattacaggagtgag  
ccaccgcgcccagccagaagaccctactttctattttggttccacatctgactgctagcatagagcctgctcccagag  
tttcataattaaaaaacaatgaatgcttctgagggactctccaagtttaggggtcagggttaggtgcaaaagggaatgatgtg  
acctgttggttttcccttttcccttgacttccaggaagctctgcccgttgggtcactgcacagccctgtctttatgtg  
gcgtagccagttagctcagtcctgctgctgagtcactagacttctagaaggaaacagactggagcaggctcctcctcagg  
ctcctccactctcctggctgcccagtgcccatcttaccattggcatggaagtccagggtggaggttgcatttccaaggg  
agttggtggctgagcacttgactccccactgtcagtttctccaggctcctggatctccaggcgagggttgaggacca  
gaggtaaccactgaagcgtgggtgtgatcactgtccccggaggtggaggccaggatagacccccatgtgacagaccag  
tgtggccaggggctcactgaccacagagcagtgaggatggccacaagtccgcctgggtctccaggaagggtgtcagga  
ctggagtgagaggcggtctgctggagacagagaggtgggctgtcaccctcagacaaagggcatttccctggataccctga  
tacaacccgtgacctctgcacccgtttgtccacactgcccctgcagaaagggtgatcccaagtgctgagtgctgctg  
tgaatacacttttggcttggctgggtgggtacgtcactctgttgtcccagtcagtggttcaaggccacctgcaagtgg  
ggataaaaggccccacttcaaagatgagaaaaactaatagagagacctgggtgagggacagcacctcagccagggtgaccaaag

FIG. 7

38/157

tgagcatcatcagcaatgggacaagctgacacagagtgccctcctgacagggcagagcagcatgtccggggccttcccacc  
cagagggcatgggtctcaggccattcaagtcaggcaaatgtggagttaggggacctcctcacaacagcaggcctgcactctg  
caagtttcaacgtcaagaatgacaaggaaagactgaggaacagtctcagactaaocggagaatgaagagacgcaacgaccc  
aatgcaatatgtgaactgtgattgtattctggaccagaaaaaaatggctacaaaagacagatttaggtccactggtaaa  
atgtgaatatagattatagcttagataaacagctcttctatcagccaggtgttagtggtcctacactgtaatcccagtaactt  
gggagggcggaggttaggtggatcacgtgaggtcaggagttcaagaccatcctggccaacatgggtgaaacccctgtctctaca  
aaaacacaaaaatttagccaggcaggatgggtgggtgcctgtaatcccagctactcaggagggtgaggcaggaggagaatcg  
cttgagccccagaggcggacattgcagtgagccaagatagcaccattgcactccagcctgggcaacacagtgagactcca  
tctcaaaaaaaaaaaaaagctcttctatcaatgctaattttcctgatcatgatcattgcattgtggatatgtaagagaatg  
atcttaggaatttagcggtaaagaggcctcatgtctgcaacttcaggaatataaatatgtaggtagatacataaaataaca  
tatgcataatgtatatagagtgtccatatatgtataagtgacattgtccacatagagtggatgtgcatacacacaagtgc  
cctgtatatatgcaagctctatatccatacatttatatgtatgtatgtgcgtgtgtgtgtctgtgtgacagcggagaaaga  
ttaagcaaatgtgccaatgttaataaaatgggaaatctaggttaaaggatatataatcattatgtcttctctgcaac  
ttttcataagtttaaaacttccaaatataaaataggaggggaactaggagggtcaaagcattgcccagggtgtgcagagctgg  
gacatgagctcaaggccacctccaggtaagtggccttgaagttcccatgcccagggaaccccaaccccttcttggggctcca  
ccatctggctcctgtcctgactgtgtccagtgccacccccacctggcctcttacgggtgactaccacgctgacagggcccc  
agcgtcctgctgccatggagcttctgcaacctcacagaagtagaagccagtatcagccctagtggccaagtgcagccggagg  
gtatgggagtgggcatcctccagcaggacatgggtcttctgtaccagctgtagcggagatcaactgggtgcctcgttgggtgt  
gttgcagactagtgtcactgtctgttctccaggatgggacctgctgggctcacctggacctcagccactgcaagggcag  
catagggagtgctgggggggtcccagccaacttccagccccagccccatacagtccccgggtctcagccaatcagcctcaa  
tctctgcacatcctacaccacccgctgctgctacaggggagcctcctggagtgtctgtcatctcttctgtcctgctcacc  
aggatccccctgaacccccaccttctgtgggggtattgtcacctgtctcatggtaaggcagggtcgggactccacacctg  
catccaggaagcactgggttatgccagtgtgtggcctgcccctgcccctcctcctccctccgtccctgagcctgagctcct  
ccctattctctggccaatacgaaccttcccaagaactcactgaagatgtggaggtgatgggggggtgagaccaaagagc  
ccacgcccgttctcagcttggcagggttagacgcccagcatcgctccagggtgctggggcagggtgcagcacaccagtcttg  
gtttggaggcgtaccccatccttgagccaactaatggaactgactgcagggttagctgctgttccactggcagggtgagtg  
gaccagctcacctggaaggtatgtcctcccgaggggctgaggaggtatctcacacccctgggggcatctgcaagtcaaca  
gtaggggggtattgggttaaggtgcttggggagggcagaggatggcacacttcttcttgcccccttttaaagctcagctcct  
aaggaagtatgcccagataaaacagcagtcctcctcaatcctcaccagggcatgctctgtctgccatgtccgtccctt  
tcccggccccctgcccctgatgcattcctatgcccataggagctgatcatgtgacgcttggccagcgtccagcctacccc  
accagctgtagttttctgcttcccattgctgtccattgctcctgctctatttgggatgagctccacacacagggtcatgg  
gtaccccaactgctagcttttagtggcctgtggaagctatttggttaagggaaccaactacctagtgggagggggccaaaggcag  
catcaaaactagcttgaaaaatagttaccagtttgaagcaagaggccaacaacacaaagaactgcatttctttaaactct  
tgttccaaagcctctctctgtgtattcaagtgttttattcttatttttttagaaagaggatctggctcagcctcagccaac  
tgagtgagagggcacaatcatagctcactgcagcctgaaactcctgggctcaagtgtatcctcctgctcagcctccga  
ggaaactgggactacagggtgcaagccaccacatccagctaatttttgggttttttaattttttttagagacagggtctcact  
atgttgcccagggtggtctcgaactccttgcccaagcaatttctcctgcttggcctcccaaagcactggggttacagg  
gtgagccactgtatccggcctcaagtgtttaatttgtgcccaggcactcttctaaatccttgacctgggtcatctccttta  
tttgtttttgttgttgttgtttttgttttgagacacagctcactctgtcaccaggctggagtgagtgagcacaat  
cacaactcaatacaactccacccccggggttcaagcaatcctagtgcctcagcctccttagtaactgggattacaggcat  
atgccaccacacccggctaatttttgtatttttagtagagacagaggtttcaccatgttgggcaagctgggtcttgaatccc  
tggcctcaagtgtatccacccgctcagcctcccaaagtgtgggattacaggcataagccacccgcccggcccatctcc  
tttaatttttatcataaatctgtgagacaggaaccatctattgttatcttctcatagactgagaaaacagagccaggca  
ggaaagggataaatccaccccccagcatcctccagcaagggttaaatccggggcacttgctgccaacctgaagctg  
catgggagctcctgttccccaaacctttgtgtctcctttctctcagcctggcctcaaggacaagcccttcgaagc  
agtaggaggttgagggaaccatttaacgaagtcttggggctcagcagtgctctcacttgccctacccctggaggccca  
ggagcagcctgggttttgcacaggagcaagggttccgtttctgtgggctggagaggggctgggtttctgtcaggagcaaca  
gatgcgctcagccacaagggtgtgtccccagacaactcacacttacttgagggtgaatctcgctctgagccctgtgatt  
ggccacggagagctggcagcgcaggatccggcctggtcctgccaggacatggcatgtggaggggtctccagggtggccga  
cgccgggtgggctcaaaactctggctgttgaagtgacagagcagcagggtcctggccttgccactgcagctctgacctgc  
tctgcagccatagctaggagtgagcaggtgaagtccacctctgtgccctcgagaagctccacggggaggcaatggt  
gggacccctgggctcctctgaggacagagacagcagtgctcaggacccgcttttgccacccctgagatccctcgctgg  
aaaccccgctgaggagagagccctggggagggggcttttaggggaaggaaagacgctttctactccagccccacttggg

FIG. 7

39/157

tgaatacaaggaggagaaccaggcctgggcctacgcggggctggaacagaggctgagactggctggggttagattcaggac  
aagggctggggctgagagccaaggggtccagaagcagcttgggaatccctcccgggggcagccaggccacccacttat  
cacctgttactgtgaccaaggtgcctttcacatctgaccagcgggttgacctcactgatctcgaagcggaagttgtaggaa  
ccagagtcctcgggctgcaggtccttcagcagcaggttgccacacctgtgctcggggttccccatgaactcgggtgcggcc  
ggggaagcgggctccaccagcttgggggtccgcccaggtggctcaccacctgcccgtggcccagtagtctgtagtaccaga  
tggccgtgatgcgctcgggcacctccacgtcggcaggggaagctgaagatgcaggggataagcagggaagaccccttcaca  
ccctgcacgtcctggggactggagacgccccatgaggcctggcctgggggaagaaaggcaggggggacagagggggaggggt  
atacaggcctcagggtgccacagagccctgcgacctgccccagagaaggtgccccagctgggctcccaaattctgcctg  
ccccagaaaatgcacacttagagcagcccttctcagtgccccaggggtcagtcacctgcccagggtgctcagaggtttg  
gtagggtggctcgaagacagatggcagcttctgcccagctcctggccactgcccagttgggcccctcccttgacctcagc  
aaccaaactatgtgacccaggcatagtctaatctgccaagggctggacttatgaaggctggacctggaagacaggcacta  
ggcccgcaagcttaacctgctgggaagaatgagggcaggaggaagcttgggcaagaagcccatagcaggttctgtgctg  
tgctcctgttgcttaaggggtgggtgcgactgcgctggctgggctcacaggggctccaggggacacctctgggacttt  
agccccagcacctgctagaagtcgagcctgtgtccccacctcctctgctggccaacccaataagagggcaggggtctta  
aagacctctgagtcagacaccagcagagagccaggaggccacgttcccagctcaggctgtgcccaggaatgcctcact  
gggtggcctgcctcagaaaagcctgtgtgtcccttggccctatccaaagtctctgctttcccagccctcaaggataccacc  
ctaaggcagatgaacagctgtttctccctcttcccacttccctgccccctcccaccacccaacccaacagggaactgga  
ggccagagatgcccagttactcactccaagacaccagctagaatgagttcttctcctgaggcttgtctcctaccacct  
gccttactaactatagaccataatggggcttactgaacttgccgaagtgtgcttttaacagtcactccctgctcaaa  
aaccttctgtggctccccattgcccgtgagatgtgaaaagtcctcatctcctgccccagctctgtccccattccctgct  
ctccgcagaccactctggggcagttctgtctgtcgaagggtccctagctgcccagctctatctccaccacagataat  
ctttgctgctgaaacctatttcaaccttcaaggagcagcatgaatttggttccagctggaacttctccttgagggtc  
ctgtagctacccagagctatctctatttctccttgccttggtttttacagcttggtgagagccatttctcaggacctaga  
actgaagatatgtgcccctagcagtgccggagcctaccaggcattcagcaaaccccttagtgactaagagaggggtgagg  
tctttaggggttcagagctgagggtcagagttggagtgagggtggcaaggcaagtctagggttgaaaggtagcatgag  
agcgtgtggaacacatacccccacaaatagagctcaatgtgtgcccagtgtaaccatatacaaaaaggcagggcctcaac  
catggagtgcccctggtcagggtgtgtcgaagggtaccatagacctgagcccaaaaggaagagatgccagaaacacata  
taagtgaactataaaactcttagaagaaaacaggtgaaaatctcatgaccttggttaggcaatgctttcttaagtat  
gaaatataaagcacgaaaaaaataaggttaagttggacttcatcaaaatttataaacttggccagacacagtggtcat  
gacctgtgaaccacacacttggggaggctgaggcaggaggatcactttagcccaggagtcaagacagagcctgggcaatac  
tgcaagactctgtctctacgaaaaatataaaacaggcctgtggtcccagctactctggaggctgaggtgggaagatcgc  
ttgagcccagggggaggggtcgaggctgcagtgagccatgattgcccactgcactccagcctgggtgacagagcaagacc  
ctgtctgaaaagaacaaacacagctgggtgcgggtggctcacgcctgtaatcccagcacttggggaggctgaggcatgca  
gatcacaaggtcaagagattgagaccatcctggccaacatgggtgaaaccccgtctctactaaaaatacaaaattagctgg  
gtgtgtgtgtgtgcacctctagtctcagctcctggggaggctgaggcaggagaatcacctgaacccaggaggcggagggtc  
acggtgagccaggtacgcacctgtactccagcctggtgacagagtgagactctctgtctcaaaaaaaaaaaaaaaaaa  
aaaaggatataaatcaaggacactatccagagagtgaaataaaaggaggacaacccacagaatgggagaaaaatatttgta  
aatcatttatctgataaaggactaatatccagaatatataaagaattctacaatgaacaacaacaaccccaaaaaaac  
catgaaatccaactccaaaaatgggcaaacacttgaataaacatttcttcaagaagatatataactggctgataagga  
catgaaaagatgctcaacatcactaggcattagggaatgcaaatcaaaacccaacacagtgagataccacttcacatcc  
gttagaatggctattcacaaaaaaacaaagcaacacagaaaaaataaatattgggtgaggatgcgaagttgaaattcttg  
tgtattgctgggtgggaatataaaatgggtccgtcactgtggaaaaaattgggtcattcctcaaaaagtcaacataggat  
taccatatgatccagcaattccactcctaggtatatatcccaaaaagtactgaaaacaggggactctaacagagtagaccaat  
gttcacggcagcactattccactaaaagggtggaacagggtcaagtgtccatcagtgaaatggatgtggataaacaactgt  
ggatgtacatacaatggaatatcaatcgcccataaggagggaatgaattctaacatatgtgaaccttgaaaacattatgc  
tcagtgaaccagccagacacaaaagggaataattgtagggttccaattacatgaaatatctagactacgtatattcag  
agactgaaagtagaataaggatagaggttaaccagggtgctgcaaggagggggagctaatgtttaatgattgctgagttctc  
agataatgaaaaagttctggaaatagtgggtgaggttacacaatatgcaaatgtacctaatgtcatgggtgtacactt  
aaagacagttaaaacagtaaaattttacattatgtatatttaccacatacacacccatgttgccaactttgcaatcctcc  
ctggctcctaaatgctgacttggccaagtgaacgaggaggctggaagtggggacaggaaactcatgacctcccagctcca  
gcccacccgctcaggggtgggtcagcagattccaatgactaccaggggtcacacctgggaagggggtgagccgagggc  
ccagggccagtcaggctgaccaggtgggacttagcctgtgcagaaggcagaagggtgcccagcagggggcacagtagcag  
ggcgggattgggacaggaaggacaccgctcccaggggacccagcctctctgcagggtgctggagtggactgatctggcc

FIG. 7

40/157

at t t a t g g a g g c c c a a g g g c t c a t c t c c a g t t c t c t a g g a a g c c c t a g g c c t c c t c c t c t t c t g g g a a g a t g c a c c c c c a  
g c c t c c a c a c c a g g t t c t t g g c c a c t g g a g a a t g a t a t a g c t g g g g c c c t g g g a c c t g g a c a c c t c a c c g t g a a g a a a g  
c a g c c t g c t g g g c a c a c t g g g g g t c a g a t g t g t c c c t g g c c a c a g g g g a t g t c a g g g t c a g c t c t g c t a t g g c c a g g g c a  
g c t a t c t t g t c c c a g c t c c c t g t t c c t c c a t t t g g g g t c c t g a a a g g g c a a t c g t g a a c c t g a t g g a a g a a t g g t g g  
g g t t c t g g a c a c a g c g a c c c t g g a a c a g g g c g c g g g g g a g g a c c c t t c c a g g a c c a c t c c c a t c a c a t a a t g t a g a g g c  
c a c c t a t g c t t a g c c c g g c c c t a a c c c c a a g g g g t c a g c c c c a c o g g a a t c c a g c c t a t t g g g t c a g c c t g t c a c c a c a  
a a g g c c a g c t t c a g c c c a g a t a a c t g t t c t g g a a c a g a a g a g c a g g g a c c g t c a g a a g g a g a t c t c t g t c c c t g t t  
t g a a a g c c t g g a g t t g a g g g g a c a g t g c c c c g c c c c c g c c c c c g c a a c t t g g g t t g c a g c t g t g g c c t a g t g a g c a c g  
c a g c g c c c c c t g g t g g t c g a g g g g g a a t t g c g g g t c c c g g g a a g g g g g c g g t g t g c c a g c a a c a g g g a g c a g g c a g c t c  
t g c a g c c c t g a a c c a t c c c t c c c t t g g g t g a c t c t t t t g g a a t c a t t g t t c c c c a g a c a g g a g g t t c c t g a g g t t c a t a  
c t t g g g t c t c c a a g t c t t g g g t g c t c t g a a g a c a g g a t t t t a a a t c c c c a c t c c t a c t a t t g g t a t g t g t t g c a t c a g g g  
t g g c t t g a g c t g g c c t g g t a c a c a g t g g g c a c t c a c g t t t g c t g c c t g t t t g a g a c c a a g t g c c t c a g g a g g t c t t g g a g  
g g t t g g c t g g g g c c c a a g t c c c t g a c c t c t g a t t c c a g a g g c c a a g t t t a g c t g g g g a a g a a g g g c a g a g g c a g t t t c  
c c t a t g g a c a g c t a g g c c c g g g t g t a g g a t t c a g t t t c t g t t c c t g a c a c c a a g g c t t c t c c c a a c t t c c c c a t t g g g  
c t a g a g a a g g a a a c a c a g g g t g a c a t g g c c a g c t g g a g g t t a c t g g c c c a c a g a t a g g g a g t c a g g g t a c g g a t g g g c  
a a t t c c t g g a g c a g a t t a t g g t c a a a a t a g t g g a a t c c c c a a t c a c a g g c c a a t g t t a a t t c t c a g a g c c a t a g a a t  
c c a t a a c t a a t g c t t a t t g g c t t t g a c a t g g g c a g t a g a a a t t t c a c a c t t c a t t c c t t a a t c t g g c t c t a a a t g c t t c t  
g g c t g g a g t g c t c a c t c t c c a a a c t g t g c t g g a c a g c a c c a g a a g c c c t c c t g t g g a c g g a c g a a g t g g t g a t g g a t g a a  
g t g g a a t t g t g c t g g g g t t a g a g t a a g g a g a t t a t c g t t g g g c t g g g c t g g g g t g a g g t c a g g g t g a t g a t c a g g t t  
g g a a t t c t g g a a g c a a a t t a a g g c t g g g a t t g g g g t g g a c a t t g a g g t t g a g t g a t g g g t g g g g t a a g g g t g a a g g t t g  
g a g t t g g g t g c a g g t g a t g g t t a a g a t a c g g t g g a g g c t g g g t t g a g a t g a g g a t g a t a g a t c g g a g t t g t g g t t a g g g  
t t g g g a t g a t g g a t g g t t g g g a t t a g a t g a g a t g a t a a a t g g t t a g g g t t g g g g t g c a a g t g a g g t g a g g g t g a g a t a a  
g g t t g g a a t a g g g c t g g g g c t g g g g c t g a g g t a g g g t c a g a g c a g g g t g a t g g t c g g g a c t g g g a t t g g g a t g c a g g t t  
g g a a g g a t t t g g g g t g g t g g a a g g g t t t c a g t t g a g t c c a t c t t g a t t a g t g t c t t g g a c t t g g g t t g g t t t g g g g t c  
c a c c a c t c g c a c c c a g a t g g a g c c c c c c c g a c c c c t g c c c c t a t c c c g c t c a g c c a g t t t c a g c c c a g c c c c g c t c c t  
a a t g c t c c a c t c a c c c t c t g g g c c a g g g a c c a g g g a c a g g g g t a c c t g c t c c a c a g a a g g a a g t g g c t g c g g c g g t g c t  
g g a c c t g c g g a g a g g a a c a g g a a g g a c g g c c a a g a g c t c c t g g t g c a g c t g g c t c c c c a g g g c t c t g c c g g c t c a a g a  
g a g a a g g a t c c c g t a t c a g g g g c t g c t t c c t c t t t c c c a a g c c t c a g c t c t a c t g t c c a a c c c a g a g g c t g g t c a g g g a  
g g c a g c t g c a g g c c t t g c g c a a t g c c a a a c g g g a a g a c c t c c a t a g g g g a a g g c c c t c g g c a a g g c c a g g g a c t t a g g  
g a c t c c a g c a a g c a g a a g t g g g a c c g c t g c a a c g c t g g a g c c t c c c c a g g c a a a g t g a g a a t g g a g t g g g g g a c t c c c a  
t t c a c c a g c c a a a t c c a c a c c c a c t c t c t c t g a g c c c c t a g g g a g g c t g g g g g a g g t g g a a g g g g g c t t c c t g c a c a c  
a g c t c t c c t c c c t g a c a c c t g a g a g g g a g g c g c g c c c a g c c t g g g g t g g g g a t g c a c t a c a c a t g c c c a g a c c a a g g c a  
g t t a t t c c a a g c c a t t c a g g a g c c t c c c t g t a c c a c t g a g t a c t c t t a a g a a c c c c a a g a g g c a g t c c c g t g t t g t g  
g g g g a c a t a a g g c c c a a t g t c c a g g a c a c t c c a g g t t t c t c t t t c c c c t c t c a c t g t c c t g t a c g t t t t t g g g t t  
t t g g c t t c t t g t t g t t g g t t t t t t c t t t t t g t t t t g t t t t g t t t t g c t t t t g a g a t g g a g t c t c a c t c t g t c g t c c a g g c t g g  
a g t g c g g t g g c a c a a t c t c a g c t c a c g g c a a c c t c c a c c t c c c g a g t t c a a g t g a t c a t c t t g c c t c a g c c t c t c g a g t a  
g c t g g g a t t a c a g g c a t g c a c c a c c a c a t c c g g c a a t t t t t g t a t t t t t g g t a g a g a c g g g t t t c a c c a t g t t g t o c a  
g g t t g g t c t t g a a c t c c t g a c c c c a a g t a a t c t g c c t g c c t c g g c c t c c c a a g t t t c t g g g a t t a c a g a g g t g a g c c g c t  
g c a c c c g g c a c t a t c c t a t a c c t t c a c c c c a c c t t g g g a c a g g a a g g a a g c c c c a c c a c a g c t a g t t a c t g t t a c  
a t t a c t g c a g c c c a t c t t a t t g a g g g c c t a g t t t g t g c a g g c a t t g a c a t g t g a a g c a g a t g t t a a t t g c t c c a a g c a a  
t c c a g a t a c a a g c a a c a a g a c t t t t t t t g g t t t t t g t t t g t t t g t t t t g t t t t t g a g a c a g t g t c t t g c t c t g  
t c g c c c a g g c t g g g g t g t c c t g g t g c a a t c t t g g c t c a c t t t g g c t c a c t g c a g c t t c g a a a t t c t a g g c t t a a g c a a t c  
c t c c t g c t t c a g c c t c c t a a g a a g c c g g g a t t a c a g g c g c t t g c c a c t a t g c a c g g c t a a t t t t t a a g t a t t t t g t a g  
a g a c g g a g t c t c c c a t g t t c c c a g g c t g g t c t c a a a c a c c t g g g c t c a a g t g a c t c t c c c a c c t c g a c c t c c c a a a g t g  
c t g g a a c t a c a g g c g t a a g c c a c c a c t c c t g g c c a t t t t t t t t t a a t t t c a c a a c t c a a a c t t a a t t c a a c t c a g t c c  
c t g c c t a c c t g t a c t g g t g g g g a g c t g g c c a c a g a a a t t a g g c c c t g t a c c c t g a g g g c a a g g c c a c g t g t g c a g g t g  
t a a a a g g a t g t g a a a c c c t a a a t c a g g g t g g a t c c a g a a t c g c a g g c c a t g g t g c c c a a a g c a g a t g t c t g g t g a c a t t  
c c a c c c t g a a a t g c t c a g g c t a c a g a g a t a t a g g t c t c t a t c a c t c t g t t c c t c t t a t a g c t c c t g t g t c c a c t c c t a  
t g c t t g g g c c a t t t c c t c t t t c g g c c a a a a c a a a a g g g t t c a t c c c a t t a c t t c c t c c c t c a a c a g c t g g t c g g a g a c  
a c c a g c t c t a g g c c t g t g g g g t t g t g a c a c a t g g g t a c c a a t c c t t c a g t c c a c t g g g a c t c t a t a c a t c c a c c c c t t g  
g c t t c a t g g t g g g a a g c a t c c c t g a c t t t g a c c t t a g c t a t a t g a c t t g c t c t g g c a a t g g a t g t g c a c g g a c a t g a  
c a c a a a c t a a g t c t t g a a a c g t a c t t a a g c a g t t t g c c t t t c c c c t g a a t t t c t g t c a t t g c c a t g c a a g g g c a g g c t  
c c a a c t a a c c t g c t g g t c c a a a g a g g a t g a c g a a c a c a t g c a g a t g a c c t g a a t c a g a c c c a t g g a t t g a a c a a a a c t c

FIG. 7

41/157

agctgagcccagcctacgtccaccagaccagtcacacctgtggatgcgtgaattttatttgctggatgctgctgagaattttg  
tggctacattagcaagatgatacaaggcctaagtccagaacaacacacccagaacttgcttacctttccttagcatgag  
gagagcaaagacttgtctaccttgattagtcagggagcactgcttctgtcatttctcttgagtatacagcaaaactaggta  
aataaataaaaaaactaggttaggtgggcacgggtggctcacgcctgtaatctcagcacttttaggagggccgaggtgggca  
gatcgcttgaggccaggagttcaagaccagcctgggtcaacgtggcgaaacctgtctctacgaaaaatacaaaaatttagc  
tgggcctgggtggcaggcgcctgtagtctcagctactcaggaggtgaggcacgagaatcgattgaacccgggaggtggag  
attgcagtgagccgagatcacactactgcactccagcctggatgacagagcgagactctgcctcaaaaatttttaaaaaa  
tgtaattttctcagtaggtcacactgttacacattcatctaataacaattattctttttcttttttttttttttttttggg  
agatagggtctcactgtcaccagggtggacagggtggagtgcaatggcacaatctcggctcactgcaactttgacctcc  
taggtcaagtgatcctcctgcctcagcctcccaagttgctgggactacaggtgagtaccaccagacgcagccaattttt  
gtatttttttgtagagatggggcttcacatggttgccagggtgggtctcacactcctgagagtcccagtcagtgatcc  
accaccttggcctcccaaagtgtctgggattacaggtatgagccaccataccagtgatatttctcgtccagatagga  
aaaccaagtcataggaggttagagaatttgccaaagacaaaacttttttggttgaaaaaaaataagttttgctacaagta  
tagaaaacaccaaataacgggtgttttaataaaaatagaagtggttccacctctcctcaagtaagtgttggcatccaaga  
tgatatgacaactccacaatcatgaaactagatcccttttatttttttggtcagtcattgtcaatgggctgcttccagt  
cattgtccaaagtggctgcttgcgtccagccactgtatctgcattcgggaaagcaggatggaggaaaggacaaagtagg  
ccatgcctctactttaagataactccttaaatgtctgatgtggtgggtcactcctgcaataccagccactcaggaggt  
gaagcaggaggttcaacttgaacccaggagcttgaggctgcagtgaactataattgtgtcactgcattccagcctgagtga  
cagagttagatcttctcttaacaacaacaaaaaaggtaattctttaattgtacactttattgctgcttatattccac  
tagaagctagaagttagtacatggtcttaagtctttattctaacagggtgatgtgccgagtgatcctgggggcccagt  
gctaagaagttagaggagatgaagccctgcctcctgtggagcaaccaggacctaattcagtcattcattccttggacc  
taccagattccagcctcaggccctgcacacctccttgccagttatctctgcagagcctcctgcctccccagaggggtgc  
cagagccctctgcttccctcagcctcagacccacttgctgtgtctcaggtggggacaaactcagtcctatagaggcccaagg  
gagatctttgagaccatctgtagtgcgttggggcagatgggagtcagttagctcactggctaaaagcctaggctcctg  
catagagctgggttcaaatccaggcatatccatcacctgctctacgatcccaggagtcactcaatgactcagaagctaga  
gtcctctggaaagcttgtatgaagagttcccaggccctggaacatcaaaagagctctgagaatgttggtctgtctgtgc  
ttctcttattgggtatctgaaggccgtactcctctctcctccacccggagtgacctctactctcatagccacacccctg  
gacctcactcattcagtcacctcaccagtcctgtcatctgtccctctggctcctcctcctgttcatgcatcttgtact  
caccacttctgtccctgagtggttcaaggctctgggggttttcagagacactgaagggaacctcctcctcaacacaacca  
caaggctctaggtggaatgaccactaagggaacgggtctctgaagccagagacttccagggaagtcaacaagcccaaggat  
gcccgttacaagaaagttaaaagacccatgtacatgtcctccgttttattccctgctcagggtctgggcatacagtgga  
acacactgcagtcaccaaggaacacgtctgtacagagtcagatggagttaagaacatttagccggctgggcgcgggtggct  
cacgcctgtagtcaggctcactcgggaggctcaggcaggagaatggcgtgaacccgggaggcggagcttgagtgacccg  
agatcggcgcgtgcactccagcctgggcccagagacagactctgtctcaaaaaaaaaaaaaaagaatatttagtctgg  
gcacggaagcttatgcctgtaatccagcactttgggaggctgaggtgggcggataacctgaggtcagaagttgagacc  
agcctggccaacatggtgaaacccatctctactaaaaatacaaaaattagccaggcatctgtaatcccagctactcagg  
aggctgaggcaggaaaatcacttgaacccaggaggtggcgggtcacagttagccaagattgtgccagtgactccagcctg  
ggtgacagagcaagactccatctaacaacacacgcacacgcacacatatttaaacaccacacccaacatctagttcaa  
gatgggtggactgagaacttgtctctgccattcctggccactccaataccactgagagcacagtaagcaagggaaggga  
gacagaagggtgggaacaggatggctgggggatgggaagtatccactgcacaggattttgatttaattctagaagatag  
aaagagggaggatcacgttcagggaacagatgcgggttaagggaacccagagccaaagcacgctgagagaaagctgcccag  
aggccgggagcagaagtggactctctacagggactcaatacacccccaaagggttggttagctggcacacgtacctctccacc  
cccacatgtaacactgcaggggcagaggaaataccctagggtgagttccagttatcaatcaacctgcctttgttcataaaga  
tgaattgggttaccaggtattaccagacaggtgagggaagacacacaaagagaagatccagaaaacaaacaggccaggcg  
cattgggtcacgcctgtaatcccagcactctgggaggccaagatgggtggatcacctgaggtcaggagtccaagacagc  
cttgccaacatggtgaaacctgtctctactaaaaatacaaaaatttagctgggtgtgggtgggttctcctgcctcagcct  
cccagtagctgggagggtgaagtaggaggttcaacttgaacccaggagctcagggtgcattgaactatgattatgtcac  
tacattccagcctgagtgacagagtgagatcttgtctcttaacaacaacaaaaaaggtaattctttaattgcaactac  
attgctgcttatattccactggctagaggttagtcacatggtctcatgtctttatttotaacaggttagtgcccagtagta  
atcctggggccagctgctaagaagttagggggagatgaagccctgcctcctgcttgaaacctagaggttagaggttgacag  
agagcggagatcatgccactgcactccagcctgggtgacagagtgagactccactctcaaaaaataaaaaataaataa  
aataaataaataaataaataaataaaaaacctacagcaagaacaaagagctattgcacccgttactgtgggcctggcagta  
ttatgacaactttttttttttttttttttttttttttttttggagatggagttttttctgtcaccagggtggagtgagtggtcaa

FIG. 7

42/157

tctcagctcactgcaacctctgcctcctgggttcaagcgattctcctgcttcagactcccaggtacctgggtattataggc  
acatcccaccacacccggctaatTTTTGTATTTTtagtagagaccgggttacaacatgggtttcaccatggtggccagtct  
agtatcaaaactcctgacctcaggtgatccacccgcctcagcctcccaaagtgccaggattacaggcatgagccaccacac  
ccatccagtggttctgacaattttatacttattaactcatttagcaacctctgagctagatgctattgttatctccactt  
acaggaaactgagtcacagaatggtacagtaaaataaccttgaccgaggtttacatggctggtaagtggcagagaaatga  
aacttgaaccaggcaacctgtttcctgagttggactctgaactactcagccatactgctattttattttattttatt  
tacttacttacttattttattgagatggggctcagccgggcatggtggcttacgctgtaatcccaccactttggggaggctg  
aggggaacagatcacttgaggccagcagttcgagaccagcctggacaacaaggagacagggtcttgctctgtcaccagg  
ctggagtgcattggtgcaatcatggttcactgcagcttcgacctccagggtcaagcgatcctcccacctcagcctccca  
agtagctgggaccataggcaccactatcatgtccagctaatacaaaaaaaaaaattatatagagatgggggtctcacta  
tattgcttagactggtctctaactcctgggcttaagcaatccaccaccttgccctcccaaagtactagaattacagggtg  
tgagccaccgtgctcagcaatacttctctttaaataaaactaaacaaaaagtgaaacaaaccaatcccgaggggaaacagat  
aattcaaggaatacaagaaaaacttatgaaaagaaagattctagtgcataatggttcataaaacaaaaaaagcagctatg  
aaaagaaagcaaaaaataattcttggaaatttaaaatagaattgctgaaataaaaagaaattcaatagaaaaggggtgtga  
acatctgtgggttttgggttccaaaattagttcactcttcttgggttaaaaaataacctgattttcctgctgtattctcaag  
cctgtgggttgggtggaattgacacctcttccatttactgcacgagagaagcatgttgctgctcacagttcattatgagg  
agccagcaggtaaagacaacactgagggcagatgaaagatatatactgacctgggtccttgggacatcattgagtcct  
tggtcaagccttacctgaagctaaagggatgagtgcttctcactgttaatagccactgatcaaatccatagaaggact  
ccaattatccttgctcagttgtgtgtccagccctcggatgagctaccatcaagggaaactgccagaccttggtgacaag  
cccaccactgaaaatggggaaaaaaccaagtccatcattgaactgaaagagctgcagttccccgaaagacaggagag  
aaggaaatgctggataaacacaacaactacttcatctcactgcacctgcaagaatccaggcacagggaagggtgggacagc  
aagtagtctttctgggttgacataacaagacaacatctgtttcttctctcagaaagaagtaagaaaggcaattggagaac  
cagaaacgacaaaaactctattccaggaccaatagaaagaaacacacatctgaagcatgattttgaacaattagtgggtgt  
aagaaaagagaatccatttgaccttgacatctgaaacagcaaatattctccatcaaggcactttagggttagaggaaatga  
tactatgtttccttctcaaaggagctatctggttacgtgggtctgacagtaaatgacgtttatacaatcataacatccgtt  
gggtttcaattttcagaatcaacctgaagataaagcatggaagatttcattataattaccgaacagcatgcaaatgttac  
aaactttgacctgaaaatgtaaaaaagagcccaggcgagtggtcagcctgtaatcctagcactttcggaggccgag  
gcggtgggtgacacaaggtcaggagtttgagaccagcttgcccaacatagtgaaccccatctctactaaaaatacaaaaca  
tagctagggcatggtggcagcgcacctgtagtcccagctacatagggaggctgaggcaggagaatcacttgaaccaggagg  
cggagtttggtgggtgagccaagatcacgcccggcactacagcctggtcaacagagcaagactccatctcaaaaaaaagaa  
aagaaaagaaaagaaaagtagacagcagaaatagagggagatggcaggtgacctgggggtgggggaggcacaattttct  
ttacatagtgatagggctcaagaaattcttctaaaaattgatgtatgaggaacgaaattttaaatatagattgttttagaa  
ctataagtagcaccaacacaggactccagtaatcacacagcaaaacaaaactgagaaaacagacatgggcccgggagatggg  
aatgggtgatttccattccctgctttaatgatgaggtgtcagtagatgctctcttgagttactaaattgagaaacaaaga  
taaagtatatgttttagaggatgattgttcaaaactccagaaaaggctgggcccgggtgggtcataacctgtaacctggcac  
tttgggaggccgaggtgggcagatcacttaagcccaggagtttgagaccagcctaggcaacatgatgaagccctatctct  
actaaaaaaaaaaaaatacaaaaaattagcaaaagcatggtgggtgcacacctgtagtcccagttattcaggatctcacta  
ctgtactccagcctgggtaatgagagtgagacctctctcaaaaaaaaaaaaaaaaaaaaaaagaagaagaagaagaa  
actccagaatatataaaaaatgaaaagagggcggaagagtggaaactgcggggtagaaaattttgcatttgtttcatctct  
ctgtatatgttgaacttttttttaacctacatgtactacgtcatatccacagcccccaaccatccaccaggcatcaact  
gctgtgttttatatgggaggttgagcaatcattcctgctccttccagttcgtcttctgtactgcagaggctagaaaact  
aaatttatatccccagattcccttccagctacctaatgctaccacctcattcagccatcattgatttccacttccac  
aaactgggtcaataattgtctcttctaaaaatattagaattgggactaagagacactacagctgaccttaaacacatgg  
gtttgaaactgtgctagtctgottatacacagatttttaaaaaataaacatattgaaaaaaattttggagagatgtgaca  
atttgaaaaaactcacaacacacatagctagaaatatcaaaaaaaactgaaacagagttaggtaggtcaggaatg  
cataaatatatattgtaattggctgtttatattatttagtagggctccagtcacagaaggctatttagtaattaagtatt  
tgagtcaaaagttatacacagatttttgatgatgaggaggttagtgcccccaacccccatgtgttcaagggttagcta  
tatttgaaatctgctggtagctggaattgggcaactaaattcacgagatgtggaacgggtcgtatacatggagaataaata  
aataaataaataaataaataaagtatttggtggcctgggtgcagtggtcataacctgtaatcccaacactttgggaggccaa  
ggcaggcagatgacttgaggtcaggagttcgagaccagcctggccaacatggtgaaataccgtctctactaaaaatacaa  
aaatcagctgggctggtggtgtgcacctgtaatccagctgcttgggaggctgaggatgagaattgcttgaaacctggg  
agtcagaggttgagtgagcagctgacacactgacccagcctgggcaagagtgaggactctgtctcaacaacaa  
caacaaaaagcaatttgttaaaagaataagaaaaacaagccacagattgggagaatatatttgcaaaacaaatatctgatg

FIG. 7



43/157

acccaaaaatacacaaagaactcttaaaacccaacactaagaatcaaacaccccatataaaaaatggacaaaacggggccag  
gcatggtggcgggtgggtcacatctgtaagtccagcactttggggagggccagcagggcagatcaccttaggtcaggagttc  
tcaactagcctggcccaacatagtgaaaccgtctctactgaaaaatacaaaaattagccgacccgtggtgacgtgcacatctgtg  
gtcccagctacttgggtgtctgcccaggaataatagcttgaaacaggggaggttgcaatgagctaagatcgccgcaactgca  
ctccagcctgggcaacacagtgagactccgtctcaaaaaaaaaaaaaatgggcaaaagatgtgaacagacacctcatcaaag  
aagatatacagatggaagataagcatatgaaaaatagcttaacctgatgtcattaaggaattacaaattgaagcaacaag  
gtaccactaaacccctattaaaaatggtcaaaatccagaatgctaaaaacaccaaattgctggtgaggatgtggagtgcag  
gactctcattcgttgggtggaaatgtaaaaatggtatggccatttttaagacagtttggcagtttcttcaaaaaactaaa  
catagtcttaccatataaatctaacaattacactcctagatagttacccaattgaggtgaaaaacttatatccagggccgag  
cacagtggtcacacctgtaatcccatcactttggggaggccaaagaggaaggattgcttgaggccaggagttctttcttt  
ttattttattttattttattttttctttttgagacagagtttctgtctgtcaccacaggctggagtgagtgagtgatctca  
tctcactgcaatctctgcctcccggttcaagcattctcctgtctcgccctctgagtagctgctgggttacaggtgca  
cgccaccatgcccagctaatttttatatttttagtagagacagggtttcaccatgtcagctagactggtcttgaactcct  
gacctcaagtgcactgctgcctcggtctcccaaatgctgggtatgctggtgaggtgagctactgtgctcagccaaggccag  
gagttccagaccagtcctggcaacacagtgatactctgtctctcaaaaaaaaaaatttttttaattagtcacatgtgggtgg  
cacacacttgtagtgtcagctactggggaggctgaggctgaggatcacttgagccccaggagtttgaggttgagtgagc  
cacgattatgcctctgcactccagcgtgggcaacagagtaagagaaaagagagagagagagagagagagagagaggaagga  
aggaaggaaggagaaaagaaaacgaaaagaaaagaaggaaggaagggaaggaagggagggaggaaggaaggaagga  
agaaagaaaaggaaggaagaaaataaaaaaggaaaagaaaagagagagagagaaaccttatatccatatacaaaacccctgcgcaa  
gaatgtttatagctgctttatttcataattgccccaaaactgggaagcaactaagaagctctccaatgggtgaaaggataaac  
aaactgtggtatctgtgcaatggagttatcattaagtaataaaaaagaagacttagcaaacacaaaaagtaatatgcata  
actaaatatgcatttagtaataatataatgcattaaagcaataaaaaaagacgtgtcaagccacaaaaagtaatatccata  
ttactaaatgaaagaagccagctctgaaaaggctacgtgctgtgtgatttcaactatttgactttctggaaaaggcaaaac  
tacagacagtaaaaagatcaaggggtgcccgggggttctggggacaggaaggatgaatagggtggagcacgggaatttttag  
gtcagtgatattattctatatgataactataatggtggatccatgatagcttttgtcaaaacccatagaaagtacaacac  
aaagagtgattcttaattgtaactatgtgcttttagtcaacaatgtattgatattagttcattaattttaacaatgtacca  
cactaatccaagatgttaataatggggaaaactgcgtgtgagggaggggatacatgggaactctgtactacagctcaataa  
ttttataaatctaaaaacttttttttaataaaagtctcttagaaaaacaataaaaaataaaaaagtaattttttttt  
tttttttttgagagggaaatcttgcctgtcaccacaggctggagtgagtgccacaatctctgctcactgcaacctctgcct  
ccttagttcaagcgattctcctgcctctacctcccagggttcaatcaattctcccacctcagcctcccagtagctgggac  
tacaggcatgcccacaccagcctggctaatgtttggtatttttagtagagatgggggtttcgccatgttggccaggctgggtc  
aagaactcctgcacctcaggtgatcgcccgctcagcctttcaaaagtgtgggttacagggttgaaccacatgcctggt  
ctaaaatgtccatttttaaaaggcagctggtggtgggcagtggtcacgcctgtaatcccagcaccatgggagaaggctg  
agcagggtgtatctcttgagcccaggagttcgaggctgtagtgtgctatgatggtgccactgcactccagcctgggtgac  
agagtgcagactctgtctctaaaaataaataaaaaatacaataaaaatttttaaaaggcaatctgcagcaagagaagatga  
agttagcaggagagcaggacagaggagagtcctgaacaacaggagaggaaggtggtcatgggtggcagctcccaggctc  
ctggcatatgttcaatccctgttcccagccctcaggaagccagatgtcccacctgcccctacatacaaaacccctggat  
ccttgacatcaactctcctttccttcatgtgctctaatgagtttttgttacttgtggaacatttaactaacatcattaa  
ccaagtggacctgtctcagagtggtgttatgagaagacagcagccaaaagacagctgcagccaagcacagtggtcatgct  
ctgtaatcctagcattttggggaggccgaggtgggtggatcacctgaggtcaggagttcgagaccagcatggccaacatgg  
cgaaacccctctccactaaaaataaaaaattatccgggtgtggtggcgagcgctatataatcccagctacttgggaggt  
tgaggcaggagaatgtcctgagcccagggggcagaggcggcagtgagccgggatcgtgccacttcaactccagcctgggtg  
aaagagcaaaactctgtctcaaaaaaaaaaaaaaaaaaagacagctgcaacaaatgtcaagttctgtgtgttttctttct  
tttctttttttctatttaatttaattttatttttagagtcagagcctccctatgtcaccacaggctggagtgagtgccacag  
tcacagctcactgtagcctcaacctcctgggtcagggcatccttccacctcagcctccttctagctgggactacaggt  
gtgtgccacgacatctggtctgtgtgttttctttcttttttttgagacggagtcttgcctgtccacccaggctggag  
gcagtgggcgcatcttggctcactgcaacctctgctcctgggtcaggaattctcctgcctccgctcctgagtagct  
gggaatcacaggcgccacaccacctgacctaatttttgtatttttagtagagacgggggtttcaccatgttggccagga  
tgggtcagatctcctgacctgtgatccactgactcggcatcccaagtgctgggttacaggcgtgagccactgcacc  
ctgctgggttgcatgttttctgacatactgtcaaaaggatactcatactaaatggcaacacattctcaagcccttct  
tttcttctctatctgctttaccacacacccatagctgctttaccgttttctccttaaaaaactcaaaaaaaccttcccc  
aacctatctatccactcttctgtcctccctcaaaacccactccattttgattctgcatacaaaaaaccaagcagagttctg  
gggaggcagaagcctcctctatgatattggaaggggggtggatgaaattgagcacacaaagccagaatcctcttttgtgg

FIG. 7



aagatgggggcagtgggcagaggggaagcaggctcatctctttccctctcttcccttcccttctcttttcacaccacacgc  
 tccgcctgggtgagctcatcgctcctcggtgtcttcaatttccacctccaggagatcgactgccaaatttctaccgcgaatc  
 ccaaactccactctgagctccagacccatctactaattgcctagtcacattttctctgtgggatcgaatcagataggat  
 tatattctgctgtgactaacagagactgaaaattcagagtcctaaacaaaagactttcttttctgtgtgctgaaaagctct  
 ggaagtacacagccctaggctggaattagggatactgtcgggggctccgcctccttgagctgcccctctgctatctcc  
 aggggtgtgtccctcaaccacactgtccaaagtgggtgactgaaggaccagctccctccactgacattccaggcaacagaatg  
 gagaagaacagggtgaggaaggaggaaccacaggcgccctccagctgtttcaaggaaggttcctggaagatgctaaataa  
 caattccgttttatatctcattagccaggcttagtgcagtgccaacacctagaagcaagaaagggtgggaaatatactgtt  
 tctgctgggttaccgtatgcatgactaaaaagagagcgctctgttcttaggaaaatgagaatagactgggttggggggag  
 ataaccagcattggctacactgcagcccccaacctcctcatgctgatgtggctgaaataaaaagagcatctcccccta  
 acatttctcttcccttagtcttgttctctttctgtaccctacacctagcagagggggcaattctgtgggtgggtcctact  
 ggcaagggggaggtctcacccaagctcagtcagggtgagcagagagaattgtacaaggaagctgaattccagctgggtct  
 ggaggaaaacactgaaaaccccaacagaagatgtccaggtggagcaaatgcagctgagagttcttgtccaagggggacaga  
 atgtcagcagagtaggacacaagggcagggtctctactgaaagcaccagggggcaagtcacccagcccttcaaccggctgg  
 ccaccaagcacatcggtccactctgctcctcctgcccgtgccaccctttgcacttttgtctgtaccatctcctccaccagg  
 agtgccttcccgctcctccatgtggtatactctcagccaccctcaggtcagatccagagccaacctcctccaagaagt  
 ctccactgtctctgggaacctcacatcacactgggagacgaagccaaaagtggggggtcagtcacagaggtgtggga  
 ggtagctcgaatcctggctggacctggtcttgggtggcatttggggacagggtattggggaggttcagcaggaaggtggga  
 cctcaaggcttctttttatgaaggaggaggaaccagctgttgacagagcagcaacagcaggggagccaccaaggctgagc  
 cttgaaacctggggccaccagggtcctgcaaggcaactgggagcagctcagcctctcacagctccatttccaggggaggaaga  
 tgatgcagcctgaggcaattggccacctggaggaagttcagctgggtggggcaggcagaggaattcctggaaggaagctc  
 accctcagcactgggttagagggaactagagacaggggtggccccaggagctgccacctcctatccccatgaataagc  
 aggtggggccgggtgaccagctgggagcaacctgggtgctcaggtaccatgctctggggcagctcctcctgagctgtgg  
 cacaggcaagagcacccggccacggggtctgcatggggcagtgacctgtctctctaagcttcagctttctcatctgtag  
 aaagaggggtgaagagtccttccccacatgggtgtgtgtgagggtaactgaagtgtagggtaggggaccagctggggcac  
 cagcacacctgaggttcagtcacaaggcccaactgggtgaagaaaacctcgccacctcctcctcctgaggcaggtg  
 gtggaggcacatgccctctccattgtgtgtgcatgagtggtgtgggtgcatgcatgtgtgtgcatgcatgtgtgtgtgcat  
 gtgtgtgcatgcatgtgtgtgcgaattgtgtgtgtgtgcgcacatgcatgtgtgttaagctctggagacaacagccagg  
 ctggggcaggagatccgggggcaagggggagctgcaaaagttctctccacagccctcagctgcataacagcactgaggtggg  
 ggatgacatgtcttagaaaacaaacttcacctctgcctaaaactgctgctgggtctgccacatcgggccaagctgtggg  
 tacttggaagaggggcctgctcctgcttccaaacccagggaagctctgaaatccccatggagggggacctggggccaggg  
 gactccccaaagcagacacagcaccacacacaggggtctaaggggacccttctccccaaaccgctgaggtaggatggatg  
 tggacaagcagctcagggtctggggctggcctagggtggaaggttcagaagcacagaaggcagcagctgcccaggggccgg  
 acatggggttgaggactagggtctggaaaaagagtttgggaggtcagccaagggcagctccacgcttggtctctctgg  
 gtcatggggagccaaggagagtagtaggagtggtgatgggtagggccagagttgtgcctttgaaagtcacctctggttgc  
 agatagagatggactgggggcagggtgagagggaaggcatctcatgagaggggaaatccctgaccacagtcacaggga  
 aacagcgggctggggagaggctgctcctggggaggacctactcagaggctgatcttctcctgctacccccaccaggg  
 agatgcaggaagctggtgaggtcagtggtggggccacaacctgctggagaggagcaggaggggagacgggagggtcacag  
 agaggtagagggcagggactctgaggacctcatccatgccacctggacaattgaatcgagcttctctagtgtggaat  
 gtgggtcactggggctatttgtatcttctctctgtcagactggatgcggcccaagggtgggtctgtgactcctc  
 atcagactgagggggcccgaatatgtgggggtgggggaaggagggagcttccaaacaggaaggtttggagggcgtggcac  
 taggacaggggatacgcccccatgagttctcacccctgctcagagcctgctgcccgtggccctcagagacgagcccatc  
 ttggagctggagtggggacaggggaaacagggtggaaggggaaaaaagaaaaaggggcaggcatgagtgtagggggccag  
 acctgcagaccccaggacatgcacagcagacacacactccaatccctcctggctctgctcagcctcaaacgaggtgcag  
 gctgcaggcccccaaccaaggcacccctgttacacacagacacacaaagacacaagatgcagcaacacacacagtaatacag  
 aaacacagtcacacactagcgccacacagcacagtcagacaaatccaagaacacactcacactgaacagacatataa  
 gcaagctcttacacatacaaaagacacacttaagcacagacacaaacacaccaacagatgcacagacggacacatagac  
 gcacactgggaggcggtctgtggagactgctgatggataccatgttggcttggtatccagaattgttggtgcagcgcac  
 tccccatgctgacctcattaaaggaggccctctggggcagtggtctgtgagaagaaaggcagcaagcctaggagggtgaca  
 ctgtagctggggagcattttggtctgggcagagaggaggaaggagagcatttcaggggagggaaaaatagcccttgcaagggt  
 ttggagtgggaagagcaaaagggtgtcccaagaactgtgctcattctgtggggctgtggggcctctctgtacggggcag  
 ccttccaggacaaaaggaggtggactgagaagcatctggaactttgtgtggggcagatggggcagggggaaggggaa  
 caggcagaacagagtcagctgcttctgaaagctccacaggggacctagcagaaattcattcattcatcgttcattcattt

FIG. 7

45/157

atttatgacaggttctctctctgtcacccaagctggagtgcgggtggtgtgatcacacttactgcagccttaaaactcctgg  
 gctcaagcaatcctccaccccagcctcctgagtaacctaggactacagacgtgagctactacacctggctaatttttaaaa  
 ctttttagagacaggggtotcctttgttgcacaggtggtctttgttgcacagccagcttcaagccatcctcccacctca  
 gactcccaaagcgctggattacaggcatgagccacgggttcccagcccagaaagagaccttccccacaaaacccctgt  
 ctggaggccaagtcccagcatcccagaagggtcccggttaggacaagaaaaaggctgtgtgcctcggaggagagggtgg  
 gtgccccctggtgcggagtgggtccccttgcggggcatggggcctcagctctgtgagactccgcgcctcctgcctgaact  
 tgagcaaagacagagctaggcctagcctggggagcccgagggcctgcctcatccgcaattctgcctctttctggccact  
 gaggaccgggttagggaagctgtaggggcagtaggggtgggaggagagagggttccgggtcaagggaactggaggggtctg  
 gggcaggggaagtggaggggtctaggggcggtcccttccccctcgcctgttgggcccctcgcctgaagggccgagggtgagcc  
 ccaaggtgggctgtggggcgggcgggggagccctcagcagctggaccgcaggaggccgaaccagggtctggaaactcc  
 tcttccccctcctccgcgcgcctctctccttctccacccctgggtgctaggggagcgggggaggcgcaggagggggc  
 tcaggagggggcccctggacgcgggaccaggctgggcccctcggcggaggcccgcgaggcaggccctcgcctgcctcg  
 cacatctgggcccggagcgcgggcccagccggcgccagggcgcggccatccggcctgggggagggggcgggcgccg  
 ggggtgggcccggcccaaggaaggcctcagggccggcgccagcttgaggggtgcgaccgcgcagcgggacgggc  
 ccaccagagctgggggcccgggatcgcgggcggtggacggggctggagctgtcgggagggcgagggtgagttctggggcg  
 gggctgcggggcgcccgagtgagaaaggcgaggaggttgcgcgtcccggtgtcggctgagacccacaaaaact  
 ccagactttgtctggtggggacagatctgcaagccctctctgcagcgtgggtgcgggctcgggaaggcacttggggcag  
 cgacttgggtctcccactgggtctcgggggaccacctcccctgtcctcgacacagctgctcagaggcttaggctggaga  
 aactcctgtgtcttaggggtcggcaggactgggtgtgggtggggtcctgctccaggggagccaaactgagggaacccggtgcc  
 taggtctgaggtggaacactgggagttgagtgatgtgtgtgtgttggggggtgcctcaggggtcccaggccctgtggtc  
 ccgaagcgcaacagagtactgagtgcccaggtgccagggtctgaggctgggacctggtcttccgggcccagcttgaggtc  
 caggtgggggtgggggtccaggttccgcctagggcagtgagcctatctgtgaggttagcgtgtgtgatgagagagactgg  
 aagggagtgatcacgaaactagggaacccctccggagagcagacgcagcgcaggaatcccaggccagaggggtaatgggggg  
 ctgtgggcagagctaggaggggtcctgtgtgttgtgtgaactggctgctccatgatgaggacccaagggggtgcggcagg  
 atgtatgggaaagcaagatgaggacagggaggggcccaggggagggcattctgtcactccctgcccagcgcaggatggg  
 gcagccacctcactcagccacacatggtcccctcttcttaagggtcctctccagaacagagtggggcagcccttgggt  
 gccgtttctgatctgtgtcttgggcccctggttggcccagctctgtggttggaaagaaacgcctgtggcctcacttagcccg  
 ctgacggagtcctgttccattacagggctggttgggtgctgagcagagagcagcaggcggaagaggcagggtcctagtga  
 ggaggggagctggaatgacctgggtgggtccttgtctacacaggcatcgtgtggagccctgaaacccagaacactgcat  
 tcccatagtgtcctcggggacatgagcgcccagctgaaacaccccccacacaggaagagatcctccaggattctgtcaggc  
 aaagggttggggcccctgcaaggtgggtcagccctgtaatcccagcacttgggaggatgaagtgggaggatcgcttgagg  
 ccagggtggtcgaggctgcagtgagccgtgatcgtgccactgcactccagcctaggcaacagagcaagaccctgtctcaaa  
 aaaaggagcgggagtgggggcagggttctgctcctaggctattgtgaagaccaggggatactctgggatgggagctctggtt  
 gtgagttcctgctggggccacaccctatcctggtcccagtgactaatcgccctgcctattgggtagaagttagggccag  
 gggctcagggtgaagctgatgactagaggtcaatatgaagccaagagttaggggtcagtcctagggccagcatatgggatca  
 ggtcggggccagagtttggcatgtgtttagccaggaaatggagttccttctgtgttaaatcaaaactgtccactcaat  
 tactctcaaacccataggcaaccacagctctgaaattagtagggccctgaagcggggctcagcctcccctcattttcagct  
 gtggactctgccagacctagcttccgggagctgggtctgtcccgtgttctgtggtcagcactctccttaagtgtccctg  
 ggtgtgtgtgtcgtcgttggcagtcagccacctcttggcttacctccacctctcggggaatccatgggatcctccttctt  
 aacacaatttttttttttttttttttggagacaggggtcactctgtcaccaggtggagtgagtggtgtgatcatg  
 acttactgcagcctcaacctccaggtccaagcaattctcccacctcagcgtccaaagtggctgggactctaggcctgca  
 ctaccacgtgggtcaatatattgtattttctgtagagatgggggttctgtgttggcagggtggtctogaactcctgagc  
 tcaagcaatccaccactttgacctccaaagtgtagaattacaggcatgagccaccacaccagcccttaacacaaat  
 tgatttagttcctacctccttcttaaagaactcatgatgactagaatgagattctcagtaagacaaaaatatctaaca  
 gaaaattcagaatctgcatcatgaagggaagaaacactgctcaaacctgcccaactcctgcttttatcaataccttgca  
 tatctggacttgttaacataagccctctgtgtcaccatctcagggaaagggcactgaggcacagagatgggcatgat  
 gactcagctaggtgccataataggctccttgcctcctgaaaccccagcataggggcactcagcatatgcacaatga  
 gagcgggtagataaactggacaagctccgatagctgggaagagtataggcccatgggttggatgtggaatccagtgtaga  
 gcgaggactcagctcttttttttttttttcttgagatgggggtcactcttctcgcccagggtggagtgcaatggcgga  
 tcttgggtcactgcaacctctgcttcccaggttcaaacgattctcctgcctcagcctcctgagtagctggaatcacaggc  
 gtctgccactacgcttggctaatttttttagtagagacaaggtttcccatgttggccagggtggtctogaatt  
 cctgacctcagggtgatccacctgcctcagcctctgaaagtgtgggattacaggagtgagccacggcgccggccagag  
 tctttttttttctgggtggttctagggtgtttctcctcagggtacaatatagcctcgggtccagaaaatccattcatatca

FIG. 7

[illegible]

**FIG. 7**

47/157

caagagggtcttgatggggcgctcctgcctgaaatgctgggtcagggtgtgtgtatgtgtgtgtgtgggggggtgggaatc  
agacttcctaattgtgtggacacccccagttcaggaactgggctatctctatgggtgtgggaggagggtgctccccctcctc  
gatcctcctcaagcagccagctgggctgtctgtggcccttgtgggcatgggagctgcagagttccctcaagtccctgatt  
agaaccacgtggcccaactttagggtggaggggagcaacccccacctcctagccagttagctctgtagctccaccttttga  
aggaggagggtctctctctcctctactcctcccaggagctcacctgcctatctccctgcccctccatctctggactg  
agtctcctgcccacacaatgcttttctgtctctctccctcccacactgtgtgtctgagggggccctagtgaggagacagt  
gagctcctggggaaagcatgggtcatctgctagggaacgttgccctctgacagcagaccctcagctcctgcccctgctc  
atgtgtgtctttcctacaccccgagctcctcctcccagcagagagccctttgcccctagcccagcaaggactgatggaga  
acctttgggtgcttggttgggggccccgctgcccatacctaccctgcttgtgcccaggatgaactgcccgtctcttctctgca  
ggctccagcccgagcggtcccagtgccctagcccacccggctccccaggagaccaggaagctgcggcattgccccct  
cacaccctcgagctccagtaagcccagagcagggagccaggtgggtgggtgacctggctgggtgtggacagggctgctgct  
ggcaagtcagtagacagggcctcagctagggaaggagggggagggggatgggggtagcaccatgctctgtccccgtacactcca  
gtcatgtgccccccagaaaggttggtcctgggtgacctggaactataaagggttaacaattctactgagaaggcgagcccc  
atgtaactaactcccaacccccccacacaggaggcaccacggccatgaagctggcctggctgaggcaagtggttccccct  
gagcatgtggcttttctaggtcccagagcactgggctccttgtggaccacagcccaccttaatacaagtcatactgaa  
ggagaggagatgtgttgccaaaaaaaaaaaaaaaaatagctctggaacccattagcttgggctgggcatgggtggctcatacct  
gtaatcccagcactttgggaggccgaggtgggcagattgctttgagctcaggagttcaagaacagcctgagcaacacagc  
gaaaccccatctctacaaaaaatacaacaattagctgggcgtgggtgggtgcacacctgtaatccagctactcgggaggct  
gagggtggagaattgcttgaacctgggaagcggaggttgagtttagccaagatcgaccactgcatttcagcctagggtga  
cagagtgagaccctgtctcaataaaaaaggaaactactagcttgaagtatccaggacaccacccctccccaccccccaagctc  
cctgggtcacccacccctggatgaaccttgcccctgggtgatccagcttagggaccccccttattccttaggggagtcagccctg  
gcataggagaaggcacacactctctggggacatgcagctgcccctctcctctctgcccacagaatcagacagctctgggttg  
tgtcactatggccacaaagagcaaggataaaaatatatacactgggtccatgtgctgtaaaaactgctggaggagcaaaa  
ttgctcccagctgggcccagactaatggctgctatgaagatgggtgtgaggtgggggtgggggctaaggcagtcagagggg  
agggtgggcagaggttggaagcaggaaagcaggagcttgagccaaacccctgcccgggctacctgagagacacgtccaag  
gctcagcctggagcctgggagggcaaggagcccgagcaggtgggcaggtgggggtggggccaggtccagggtggtctgga  
ccacagtcagtgaggggagtgctctccccatgcagaggaagcattggggctgtgggggatgggggtgtccctgctggct  
gcttgctaattctaagtcttgcaactgcagaaacccgaggtccgagccccccagcagccctccttctctgtgtgggtggc  
cattgacttcggcaccacgtctagtggctatgctttcagctttgcccagtgacccctgagggccatccacatgatgaggtgag  
gtcggctgggctgagagagtgaggtggggagggtggggagttcctcataccttggtccaaaagtactgtcacccgagaca  
tggggctccttgagggccctgccacccccagctctggggctcccactggggtaaaagtggaggatggggcccgagccct  
gggtccttgccctcaactggagcagctcccacccctctgtaggaaaagtgcagactttgttttttaatttttgaatttttttg  
agatataacaaacataacaataaaaatatgcaaatgttttttattttccttttttttttttggggcgagctcactctg  
tcgctaggctggagtgcaatggcagatcttagctcattgcacgctccacccctcctcgggttcaagcgattctcctgcctc  
agcctcccaagtagctgggaatacaggtgtgcccactatgcctggctaacttttgtgggggttttttgttttttttttgag  
acggagctcactctgttgtccaggtggagtgagtgacgcaatctcagctcactgcaagctctgcccctgggttcat  
gccattctcctgcccagcctcttgagtactgggactacaggcgacccgcccacacccagctaattttttttttttttt  
tgtatttttagtagagacgggtttcacccgtgttagccaggctggtctcaaatcctgacttcgggtgatccgcccagcct  
cggcctcccagagtgctaggattacagatgtgagccaccacacctggccttttttttttttttcttttgagacagggctcctg  
ctctgtcacacaagctggagtgagtggtgtgatcatagctcactgcagcctcaaaactcctgggtcaagcaatcctccc  
acttcagcctcctgagtagctaggactacagacagcaccaccacacctggctaatttaaaaaaaaaaaaaaattttttttt  
ggagagacagggctcactgtgttgccagggtggtctgaattcctggcctcaagtgattatcccacttttagcatccca  
aaaatgctggaattgcaggcgtgagccaccatgcccagcccacaaatatttttctaagcttgtagacaaacacacacata  
tacaacacatatatgtgtttttttaagtaaaagttagatatacaacttgctttgttttaagcttcaacactattttgtag  
atatttcatgtcagcacatacgaagctaccactttctttgaacggcctagattccatagtagacagatgtgtcataattta  
tccattccctatttgacagtttttaggggggttttttctcctctcattttttcactattacaaaaagtgagcagtaatat  
ccttctttgaaaacttgtgcgcaatctccatggggtagattcctagaggtgattcctaggtccaacatctcaaatcttt  
tttcttttcttttttagtaaaaaggaaaagaaaaataataaaacaagaaataaaaaataaaaacaagaaaatgaaggtt  
ctaagggtgagttgacaagtacacagttgttttgaaagacatgtttcaaaaatccattcaaatgttgggttttatgtg  
ggacctcaaggacctgggacctcaagaccccagtcctctgaatatgtcctagagggtcggcaggcagtaggtgaagcta  
caggagacctcaagctctcaaccagactgggatgggaggtgaaggggcttcccaggcctggagactggatgggaagg  
gtgaatggagacatagacagacaggttaactgctcccaaaatacatgggacttgagcttcatgagactggactgctcctcc  
ttccacatcacaggaaaccacaaaggctctgagcacttacagccctctaggagctgcagggcacccagatgtcttgctg

FIG. 7

48/157

atcaacacacatggctgaggccatgctggggctaaggacacatccaggaacactgtctctgtcctgtggaggtgacagt  
acaggaaataactgggggacattccagcctcctcccaagctctctctccaccagacaccaagcccctaacatctctag  
aatccccatctacttctctccatggccactgatcgtgcccaagttcagccccgaccttctctcacctggatggctgcaat  
atcctcctccttgggttctctctccatgctcccttccctgcattctgttgctccatgggagccagtgaccttctcttct  
aaaggcfaatgtggctcatcctcttggtactttaaatccacgaagacagaattctgaatggccttctgtgctcttcatg  
atctcaccttagcattcttgtgttttctcctgctgctatggagccgacccatagctagtctcctcctgctctttcacatc  
tgtaagtttttgacatgctgttccctctgctggattcccaaccaagaaggagggttgactgactgccctgtgctc  
cgcaggaaatgggagggcgagacccgggctggcccaaccagaagaccccgacctgctgctgactccggagggcgc  
cttccacagctttggctacacccgcccgcgattactaccatgacctggaccccgagaggcgaggactggctctacttgc  
agaagtccaagatgaagatccacagcgccacgggtgagtcacagggctccagacagggaggcggggcccagcatggaaaagg  
gcagggtcaatgggggtgggtgggacaaaacaaacgtgtgaggacgggcccgatggagtgcgtggctgagagggggcg  
ggctaaaggagagcgtcggactccggtgtgggaggagctcagaaatgaggtggaggcggggctaagtgtgggtgggctaa  
tagtgaagttgggttgacagggggtgggctaaaggaggggtcggggcagagctaagtgcacatggggcaagagtgg  
gacggtggtaagaggaggggaagctccaggaaaacgggggtgatttaagagcgaggtcgtcaggaaaatgagtgccaaagct  
gaggcctcctgcagagccgcccgtgtgtccctgccaggatctcaccttgaagaccagctagaggcagtaaatggaaagac  
gatgcccgcctggaggtgttcgcccagcctgcgcttcttcaggagagcagcccttcagggtgcgctgcggccccacct  
ctgcccactgtggcagggaacccctattttccctcatccgaaacggctccccatcccgtccccgacattggatgggtg  
gcccgcgggagctcagaggtcatcttccagtaacctcctccttttggctgtagagcctgcaccaagccatact  
gatgggagggggggcgattcttccagctctgctgggaagtccttctgtgatttgatttagtaacctccagttccgcagagg  
gctgaagaccacccctcctccaagccagcttctcctcactgccccctcctgtaccaggagctgagggagcagagcccat  
cgctgccagagaaggacactgtgcgctgggtgttgacgggtgcctgccatctggaaacagccagccaagcagttcatgcgg  
gaggctgcctacctgggtgaggacgtgcaggcgggcccgagaacactgctcagggaaggccaggcctgtccccatgcttgc  
atgcacccacccaccccttgagaccacagagtcattgtggaaagaacttcagcctgctcctgatgggagtttgtagagttg  
ctcccaccagaagagggtgggctgcaggaaacaggggacagaggggacaaaagacacagcccagggcaggtgtagagct  
gccttaagttaggtgtcctaaaagcagagcccaagatggagagtccttttttttttttttttttttttttttttttttgagacagctctt  
gctctgtgcgccaggctggagcgagtggtgctgctcggctcactgcaagctctgcctcccagggttcacgccattctcc  
tgcctcagcctcccagtagctgggactacaggtgccccgccaccacgcccggctaatttttttgatttttagtagagacg  
gggtttcatcgtgttagccaggatgggttcaatctcctgaccttgtgatccgcccgcctcggcctcccaaagtgctggga  
ttacaggcgtgagccaccgcgcggcccaagatggagagtccttatatgtgaggtctattgaagaagggtttctcaggagaa  
aggcgaggggaggggagggacaggagagaaggagctatgaaggatgtggtctttgctggagtctagcctcaggttttagcct  
cagctagatcccatggggagctgcagagggagaactgtagcacagagttggggccggcttcttgacacatccgtatccgtc  
tgtcactggttctggggaaatgggagggaaacctctccaagttagggcattcccatcagctgagggcagttatccagag  
gaggttagcagctgtgagccaatagcagccaacactcacagtggcaggcagggcaccacagaacattcacttccattggcat  
cagtttttgagggaacccccaccaccccccactttgttgggttcttgagcaggggaaataggactatgaaataaaggg  
aggcgattctagccaatcagctacaggtctctaccactgagatgctcaaaggcttagagcaaggagtggggggggctc  
ccgtggagacagagagggcatagcctcaagactcattgcttggctcgttaacccttagagccttggatacacagacccaaagctg  
tccaaaagggtctcaagcctgctgtggacacccctggagggaacccccaatcctcccatacctatgggcccaggcatgacc  
cagggccaggtctgtaaatgtctacacttttctactggaaggagctaaccctgagcacagtggaacatccctgtgcaagat  
caggtatggaagcaaggctgtccccagccagaacaaacacctgctccccagcgtcccttgcctattttccaccctcct  
tctcacgctgttccctggctgctcaggaaagtctggcctgggtaaaagtctccattcacacccttatccctctcccta  
tctggtcagaatctggggaacccctacaaaaccataataacttcccacagggtacccctaccctgaaggaaataactcca  
gatctcaagtgttttctccaccaactggcctgtcttggtgctggaatttcagcttccctcctgctgtatcagattcc  
ctgggaacaaagtctcctggaagtgggtggcctatggctaccgtctgacaccagcctagtggagaaagaggggttcaat  
gagctaaaagggtccccacacccctcctctgagccatcacacacccccacattgtgctagagctctctgccaataccccca  
gccagcgtcagctggcgccttgtaacctgtgaaggggaaccttgcctgagccccgctatctgggaaatttgagggggca  
gatccagggggttctaggtctgcatctgtctagagcttttctgctgggcaacttcttgcatatttgggcaggggcca  
ttccctccccagcctggctcacagcgtggctatgaggcagctccaaatttgctgagcacagaggggctgagagccctaac  
atgggtgggggtgtgatggaggaggaggtgggagccccacaggcccggttccccacctcacagtgccatcttaggtgtga  
cagcccacagtgctgctgacctgcccaccacccatccccaggctggactagtgtcccagagagaatgcagagcagctac  
tcacgcccctggagcccaggccgctcgggtatactgccgcaagctgcgctgcaaccagctcctggacctgagtgccgg  
gccccagggtgggtgggcccctgggtgagcgcgctccatcgactccagcttccgtcagggtgagctgccccggggacacc  
accacccctggagggtcagagggctcactgaagccagaagctcagccatgtctagatgaaggggagaggggtaccaccc  
tggaggagcccaatttgagcaggcagaaacatggctgggtgagcctgtctgaggaaggaggacagccctgcccagggt

FIG. 7

49/157

caacctcgctcaggggtgggacatgcctacctaaggccagcccaatctcaggggtggagatttcaccttgccctaggggag  
caaagtctgaggaggggcaggactccgcccagccctgagctctgagggttgggggagacacagactaccccaggaggaaact  
gctgggggtactgacaaaggaggaagtccagccactgcatcagctgtcccttcactccctcatcttccccgacacacat  
cagccccaagctcctgctgccaggaccaggcaccagggccaaggacagctatggctcatcctcccaaatgccatctccag  
ggcaggggagaagccctaagtccctgagctccctctgagactccaaagacctacctgcctccgctcctccaaacccctctaa  
ccctgattttgccaatgacctgagacctcctgcgttaaagggaaggccctgtgtccataaatatttccccacagctgttggga  
tacaggggtgggagtttgggggttcaggattgcctctccagctcaggagcaggttggagtttcaggagcactggctgctcc  
cagtgcccatggaggtcctgggcaggaggatgggagttgaacgccatagctggagcacctccttctaactcactccctg  
ctgtctcctgacccccagctcgggagcagctgcgaaggtcccgccacagccgcacgttccctggtaggtcaggcgttagga  
gagctgtgggagagatgcaagcaggtagggggaaaggggacggagtggtatccttggccctacccgggcacacataac  
tgatgggggggaagggcatgtttgcaaagccctctcttccctccattccgtgtaccaacctggccgtccctccacag  
tcacccgcacccccacccactcacagcggcgccctaaactcccactcctccaggggattctccgcggacgctcgggtgg  
agttgcagagcctctggaaccatttctccccacacccctgcgcccataatgtgggtgtctgaggttcaagcacctgaagcc  
cctcacgtccctccccgacccctgcagacaggccttgggacccggggcagggctggaggctgggcgaggtggagggggc  
gcagggctgaggggtgcagggcgcgccacagagtggtgtgcccgcgtcgcgcgcgcaggagaccgctacgtgggtggccgact  
ggggcggaggcaccgtggacctgacggtgcaccagctggagcagcccatggcacccctcaaggagctctacaaggcatct  
ggtagtagccaggcggcgccccggtacccagcgcgacccgggtccggccccgcactgccccctggcgcccggcag  
cgctgacgcccctcttcgccccctgctccacccccaggggccccttatggcgcggtgggctggacctggccttcgagcagc  
tgctgtgcccgcactcttcggcgaggacttcacgcccacttcaaaaggcaacggccggcagcctgggtagatctgacctc  
gccttcgaggctcgcaagcgcactgctggcccacaccgtgcagggcgctcaacatctcgctgcccttctccttcattga  
cttctaccgcaagcagcggggccacaacgtggagaccgctctgcgcaggagcaggtgggtcctgagcccgcgggctcagg  
caggggttgcgcaccgggaatgaccgtgcactggagggtccggggcccaaggaaacgggtgggggtctgcctgattcatc  
ccacatatacactaagccagcagggcgctcgggggtggggcgccgggagcgcgagtgagtgccccagccagcagggctcc  
acccacggaatccgcagcccgaactggggcaagacagagaatcatagcggggaggcggcaatgcctatctcctccagcc  
ttctctacacccccacccgggcccctgcgggcccctgctcctcggttccctgcaccaaagcaaggggaggccctccca  
ggacctcgtaacctggaacctggagcaggtggcaactaaatcctctgagttagtaggggtggagataagggaactaacatcc  
cgaggtccagtcctccagacaccagctgcagtcgggtgccaggcacttctgcctggaggcagaggttagagaataaggac  
cacggacccccaaactggggcaagcagctgggcccctgaccgatggatatttggcccttccaccaccaacagcgtgaacttc  
gtgaagtgggtcctcacaggggatgctccgaatgtcttgtgaagccatgaacgagctcttccagccaccgtcagcgggat  
catccagcacataggtgagcacctgagcttgggtccccaccccgcccctacatgaacaaacagatgcagaataattcccc  
ctatcagtgccctagatacctccacacatccatacactgtgatgagacctagaatcatctagaacacctgcgggatgaagt  
gcagtgggtgattaagagctagagggttgatattgagcttgcgaaggcaaaaaacttctgggtgcctcagtttccccattt  
ataaaatgggggtgatagattgggttctcaaaacgltatcacagggataaaaatgagctgaagtacctagagtgcacaa  
tgtcttgcacacaatgtctaggtgttaatacgtgttaaaatgcataccttatctctcgctccacagctcgtgtgggag  
agaagtggggagcgtgagtggttggggagggcgaagccctcgaggactcccgtagctctcaaagaaagtgcctcaaatggct  
acttcttagtcgcccaggtaggtacaggctaggagggggagggcgccggtggccgcctagtgggtggcctcagtggtctctc  
tcccccgccccttctcctctgcccccttcccccgcgtcccccgctcctgtcccgagaggccctgctggcacggccggag  
gtgcaggggtgtaagctgctgttccctagtgggcggcttcgcgcagtcagcgggtgctgcagcacgcgggtgcaggcgccgt  
ggggcgcccggtctgcgtgtcgtggtcccgacagcgtgggctcaccatcctcaaaggcgcggtgctgttcggccagg  
cgccggcggtggtgcgggtccgcgcgtcgcgcgtcacctatggcggtggcggtgctcaacgccttctgtcctggggcgccac  
ccgcccgaaaagctgctggttcgcgcagggccgcgcgtggtgcacggcagctcttcgagcgcttcgtggccgcgagcagtc  
gggtggccctggggcagggaggtgcggcgagctactgcccggcgctccggccagcggcgctactcatcaacctgtact  
gctgcgcggcagaggatgcgcgcttcatcacgacccccggcggtgcgcaaatgtggcgcgctcagcctcgagcttgagccc  
gcgcactgcggccaggacaccgcggcgcgccctccggcgccgcgcgagatccgcgcgcgcctgcagtttggcgacaccga  
aattaaaggtcaccccgctcgagctcagcaccaatcgctccgtgcgcgcgtccatcgacttcttcttccaaactgaggcgcg  
ccggcgcggtgccagcgccgtctgcgcggccccgcctcttccggttcaggggcctgcggagcggggttggggcgggggaa  
acgatagttctgcagctctgcgccttccacgcctccagccccgggggagataagggtcatgggagagtggttggggacac  
acccagagactggccttggggatggggcactggtccgctgactgccagggtgaagggaacccgccaaggactgaacgggtaa  
gagaagaggtttgcaagacagagcgcgcagcccggaaggggcatgtgaccccggaaggaaagcgaacagaagagtcct  
ggtctgaacttggccgagtaggggtgggggtgggatggcaggaggagccgcaggagggaaggaggttgtgcagggtctgga  
cctgcagggtgaagtctactcatcgaccgactcagcccaacccgggagccaggcagaaaaacccctgtgcctgaggaag  
tgactggaagtggactccagaggggacaggtgtgtgcaacagctcctgggtgggtgctgaccacccaaatagactgtgaa  
ttgtggaaagggcagtagatctctaatgtggaggtgggaacattattgtgggtggaggcaattatgagggttagcatttctt

FIG. 7

50/157

togagacaaaacacccgtctgggaaggccccaaggtcagcttatgaaggaccccacttgacccccacccagccatggaa  
gagcagctggagggtggatggggaggccagagggagcaatgaggggtgggtcccagctctgctattgactcggtatgcctt  
taggacattctcttaccgctcatgggcctcagtttctctaaagtgtgaaatgtcaggcacttccctctaactggcatgcaa  
cagccccacctgctgagagccctgaggtgacaataaaacatttatgctcaaggggaagccacagcctgctgatatggcg  
tggagacccataatagtgggaggaatgcaaggggtcccggtgctagagagagaagggagaaaagcttccagctgtgcatagg  
gaactgaccagaaggggtgctgctgtctcccatcaagcatcccaacaactccactgcttaagacctctctggcctaca  
catgaggtccctctctcctcatccaattaatgtcttgggaagccagcttctggcctaaaatgccaccaacctgtgcatac  
ctcttgtggggctaggtgctataataaccacgcggtgccccctgctcctgagtgagtctaccaagctcttccctggccca  
totgcaaaggagtaggcattacccccacccagagaaacaaaaatccacctggcctccggtatccactggaagtattttc  
tttagggttctatcccaaccagtcgcttaaaaaccaagtaacacagacctgaggggtgggggctggggactgcacctccc  
tctactcatggtggacagcagtgagggactaggagggggcaggagaggtggctgaagcaaggcagcagtaatggggccac  
gacgcccacagagccagctccgtcctctcccagacctgggtgggagtcctgtgggttgggggtggggagtgggggaccac  
cccaggccctccctctccctcctcagacagcctccttccgggtcaacccattcttccggcaggagagctgaggcacac  
agagaggaggaagtgggagaggaggaagggggaggggcagggtggcagcacaatgaaggcagaggtgagagggctgggc  
aaggccactccacccccacccccacccagagagggggcaggaagccacaccatcacgcagcatgtcggggggaacaaggc  
gggggttaaggctgaggggcccggggcaggcggggcctcgggcctcagtcaaaagccgtgccagtcgctgtgctctgagtc  
gtattccagctcggcgccacacacttgacaccatccagcagcatgggcgtgcgtgggtgcccgtctgcaaggcaggggtg  
caagtcagtgccatgctggcccccgcccccacccatgcggccactaagggggacccctccctccctcagggatcagct  
ggaggtagggacctgccaaggaggttgagaacccctgagccgggcaaggatcccttggttcagccttggttccctgaggag  
gacagaaacccctcgcagtcgagcttgatccctcctccaaccaggagcctcacgctctcaccatgacgcgggctgc  
accgtcacgcgcacacaggtgccagtgccaccttcgcaggctagcagtatctcctctttgaggacgcccctcttctgtgcgc  
cgagtactcacacttgacactataacctgcccggggacaggggcaatttggtcagcacctccccagcctccccgatcctg  
ccctggcactcacgggcccctacccaaccaactggggggactgcatgctaagccccccaagaaagggtgagggatggc  
ccttctggagcccagagaccattccccctagcagggtgacaggtctcaaaaaaccattcttcagtgcgctggacatg  
cctccagcctgtgggcccaccccaatgtggctccaagagtgaaatgaagttagggtccaattcagggttccaaaagaaggctct  
ggctgttttctccaaaagggaaggcagggagaggcggtgatgaggagtgaggggggagggcagggtaggcttgagcag  
atccgatggcaagaggttaaggcctgaggggtgtcaatccattgaagggcaaggctcatgaagaggctggtgacggggaca  
tgaaactgaagctggagctccacgaaactgacacctagcctctgccaggctgggagtggggcatggggcagggcctgaa  
gtgagcctgataggaacagcagcctgaaacccaaggtcgggactgggtgggtgctaggagttatccacacggtgtgtgta  
gctccagcaggttttctagagggttagggagggcagggaggaaggctggaggcttcaaacaggttccctcagcagctcca  
tottggttactgccccacggaggttaacctcacaccatgggcaggtacagggaggtatgagctcatggacttctgttcag  
gacagggaggaaggcctggaggtgtgggacctgcatgctgccacagtgcaagctcacaaagaggtgccacatccccgact  
tgctgagctgcccattccacctagttaggaagtgaaggagcaccctatgttctcactcccacacccatccaggccctgctg  
gaggaggggacgcaccttcagggaacgggcaccaacgctgaggagcttgagggtgaggctggggacaggtgctctgcggaca  
tcttgctcagcctgtgcactggggcagagtgaaagttatctcataacctgtgcaggatcttcaggaaagccaacctacag  
caggagaggtgagggagggagggcaccttcaactcctgcttcccacaggccccaccagttgtgccacctatgatcccatgc  
gccccagcacctgcttccccagagagcccagaaagcccaactccctccatcactccagctgatcgacccccagccctttc  
acagacccttgagggtcccagccctacagcttggccagcaagatcctcagcatctgtctcaggggccccaaagtctca  
gtggctcatttcatagacaggaaaatggaaccacaaaatgactggccaaaggctatgtagtgtggccagggtcaagtc  
tccctagacttcttgccattcctgctatgcatggcacaagggcaccacccactgctacaagggaaagttcaaag  
tcccatgggtcattccactgtccccatcactgccagtatcccagggtcagggtgctgggtcctccattcgctcataggat  
gacaccaggccacagatagcgtgggtgagggatgggatcaaattagagatatgacagtagaccagctgtgtcagcccag  
agctggccacagacctggcaccagccactgaactggggctcctgggtgcccctgctggcactggagctggggatgcctgcc  
tgcccactcaaaggctgtgtccaatggccatactcacgacccaagcaatttctgttgttcatccacagtgcttggtgac  
aaccaccctctgggacaaatgccattcttggaactccagtagtatgaagggtcacaagccaggggtggttgctgagcagg  
ggggctggtgggggtgcccagccagcaggcaagggtgctgctgactccagcttgcatgggcaacagagctcctgct  
cacagttactgggggtgggcagccccatccctggggggccaaactgggactggctgcagagagtttagacctcaatggga  
ccaggttgatattgctacatgacaaacttcaatccatgtgccccctcatgctctgctggtagggggcacatccctcaag  
aaggcttctggctgtacctgtgaactgtgacctgtgacccaggggtactacttatagaattttccaggggaaatagag  
atgtgcagaaaaatactctcgctgcccattttataacaatgaaaaacaaaacagccgtaacacccaggaacaggggagcc  
agctaagtcaatgatgacacatatcatggaccagtaggcagttgtaaaatctcgtggaaaattatttactgacacaggtg  
ggagacagttcacaaacagaagcaagcaaaaaccaggtgtggacacggtagggccagctttgtttttggaaacaaaagcttt  
ttaaataaagctggatccatcaccatagtggtttcagaaaaacaaataaataaataaataacattttaaaaagctggaat

FIG. 7



51/157

aagccaggtgtggtgggtcacgcctgtaatccagcactttgggaggccgacacggacggatcacgaggtcaggagatcg  
agaccatcctggccaacagagtgaaaccccgctctctactaaaaatataaaaaattagctgagagtggtggcacgcgcctg  
taatccagctactctggaggctgaggcaggagaattgcttgaacctgggagctggaggttgagtgagctgagatggcg  
ccactgtactccagcctgggtgacagcaagactccgtctcaaaaagaaaaaaaagctgaaataacacacatcagaatat  
taaccatggttattttgggtagtatgactgaccgtaacgtgaagaggatgcgatattgactgtgctttccttctttgctg  
gtggatataaaaaactggaatggacattgctatgacgggatggagctttcttacaagctgagcatgagcttaccatata  
atctagcaatttctactcctagatgactacccaagtgaagtgaacttatatccagggccgagcacagtggtcacgcct  
gtaatccagcactttgggaggccgacacggacggatcacgaggtcaggagatcgagaccatcctggccaacagagtgaa  
accccgctctctactaaaaatataaaaaatttagctgagagtggtggcacgcgcctgtaatccagctactctggaggctga  
ggcaggagaattgcttgaacctgggagctggaggttgagtgagctgagatggcgccactgtactccagcctgggtgacag  
caagactccgtctcaaaaagaaaaaaaagctgaaataacacacatcagaatattaaccatggttattttgggtagta  
tgactgtgggtacttttaatttcttcatattttctctgtcttccaaattttctcatgtctatttaagacatgagtggtat  
ttgcaatgaggggagagagctatttttaagtgttgggttgggtccatttactctcttgggggaagctgtcagctgtaggaca  
aagagtgggaacttcagagttggacagaaatggatacaaacctgggccccccagggttcttttttggttgttgggttgaga  
cagagtccttctgttgccaggctggagtgagtgagggacaaatctcagctcactgcaacctcagcctcctgggttcaagca  
attcttctgcctcagcctcccaagtagctgggactacaggcgtgcccaccacgctcagctaaccttttgtatttttagta  
gagacgggtttcgacccattggccaggctaaacctgggcccccttttgagcagcgcagcagacccccactcagggccatc  
tcaatgggcccagacctccccgacccaagggcactcctgttcaactctggacccctgatttattgaccaacctgaaggcc  
tcagtcctccattttccccaatccagctccaccaagcagaaaacaggggtgatataactacttcaagagcctaactgag  
ccagagatttggaagaagggttaaaaaaaagtgtgcacctgttattgccaatagtttagcttcacacctgtatcacata  
catacatttctccatttctcacaccaactctaggggatgtgattgtctccacttgagagaaaaagaaactcaggtgaccttc  
ccaaggctcacagagccagaatggctggcctagacctgaactcaggccactggcaccacagccacgacactgccttccatt  
gctattgtctgggtgaagcagatgacagcagtggtctctgctcatcttgatctggatgacattctaaatgctctcatctca  
ttgaaatctcactgcctcccatctgacagatggggcaaccaaggcagggaaaggcacagacatgcctataagccccgc  
tgagagggaacgggttaggcaaccagaagccaggcagagccggggaggggaagaagctgctgatggggtaggtgtgcatg  
taaccttgaccagaaagctgctgtcactctcctgggtgacctgaccaccagctcatgagcttctcatcaaagtggacgt  
ggctgtgggatccttctgcatcgtggcctgcccgaaggcgatactccggactctgggctgttgccctaggaagagtggg  
agagggctctgagggctttccctgtccccaggaaactcctccccctgtccccctcgtcccccccaagactgccttgaca  
tcatccagctcccactggctgggctctttcagggttagatggacactaggatcatgaggcttgaggcctccccgcggc  
cccagactgccctcccccaaaaatccatccccctgagatgagcgtggggggactcagggaactggctaataatgaccttg  
cttgagggggtggggctgggtgccagagccaggaagggaacagtcttccagactcccaccaagccaggggcagcgtgggact  
cagcccagtcctctggataacctgcccagtgctctccttcacagtgcaagctcccatccctggggcctaccctaaagc  
tgtgtcagtgcatgagggccctgcctgccatcctcatgacccaaacctgaaaggccagggaacagggtctgggg  
gatctgctgggtggcagaaacaggggggatcaaaaacacactcaggccacctggcgaggctgcagctgccactaaa  
cctcactggccctggcagcaagaagagagaatccaagggaagagcttccagaccacctcaagaccccccttcttctctgt  
ctagcccccatcaaggccctcagctgtctctgagctgggtgggttctccaaggccagaaagggcagacggatgagaca  
ggcagggtagggcgaggccgagctgaaggcctggactggatgaatgctcaggtgttagaggccgagagcacggccctagt  
cctctctgacccctggccctaggccctcccaccaagagccctgcccagggtcctcttggcgggaggacctgatccagct  
ttgagccgctcttggatactcggaggccaggggaccagcctggcacatccacaatccctggtagcccccacagaacc  
ctcaaatctcaccagatcaccctggcctaaaacctccttctccctctggactctgaaagaaacccaaactccttcc  
acagcctgtctggccctcccacatcaccagctctcagtgctcaccataagccctcggctcccccaacccccacagtggg  
acggcagggaacaaagagcacacacctccacagaccagcctaaggagtcagcagggtctctgaggggtcctgctagggt  
tgctgagccccagaacctggcacaacactgacatagtcacatcacattcgcacaccagcacacacgactcacacat  
gcacacacacacacaggtggtgccccctacctgcagcaatggccctcctaccaggggccacccggaagagcttgtgcaa  
gtcctaccacgcccaggaaggcacttaaccttggtggctgcagccatggaagccctcctgccacgagctcctgccagac  
accgccactcacgctcagcagcctccatgctcaggggacaccagcgggagcctgaaggataggagtggggagggcagg  
ggtcaggggccacctgggacccatgggtcagggacaccagagagctcctttaggaaggactcaaacccctaccgctct  
cagcagcccaagtgggtgctgactctctagcacgggggttctgcccacctctaccgactgccacattcaccaagag  
ccctcacccttctctgataccccagtgactcagggtcctgcccagcatgtgctgagctgcagcttctagggtgccaagcaa  
cagcataggactcctatccccagcacccccagagactggcaggagggggcagcctggaaaggaggactttattgggtattt  
ccagggaaccagttctggtgctgctgaccccaaaggctgggctggagctacctattcaggtcacacaagcaatgcagct  
gggtgcggtcagaggcaacagggtgctcgatgggtgggcaagaacaggggaccacataaagtaagcgtgccctagagctt  
tccctcctgggtgatcggtcagggccatattgcaaaaagatgcctgtggaggggtgtgccccacctgaacctcattcag

FIG. 7



52/157

accctgcctcagtgctcaggcctgcctccgctcagctccccaacggcagcctcagcacaggggcagtgagaggcaccoc  
aggaaagcctccaatgggcccagctgggaccatctgagcatcaaaaagaataatgagtgcaactgatggaaacagatggaa  
cacataaaagcagtgcatcacagagattattagaaaagaagagagaaagcaaaacaaacaaacccaagccctcaaaagc  
cctcatttgcaccactgacggtagcagtgccctctttactctgaaagctgaggattaaagaggaagaattcagcctgtc  
tcttggcctttcttgttaaccaaattgccttggtggttgatggaaagctcttctttcggaagaattcttgctaataaata  
acaaagaatgaccaaattgctaagtcattctgcaacccctaaaggaaacaaatgctggaggcaacaagaaccagtggtgc  
taaaccagaggggaaggttgacaggaagcagatactcacaggcgcccaagcgtcactgacagaggacctgggtatggttt  
tggtttttttttttctgagacggagtttcaactctgtcaccaggctgggtgcaatggcagatctcggttaactgc  
aacatctgcctcctgggttcgagcgactctccgcctcagcatcccagtagctgggactacaggcgcccgccaccacgc  
ctagctaatttttgatttttttagtagagacagggtttcaccatggtggccaggctggtctcgaactcctgacctcaagca  
atctgctcacttcggcctcccaaagtgttgggttacagggtgtgagccaccatgcctggccgaggacctgggtgtttaaga  
ggaatggcacaagtttgcaatggaggtatgtgacttctccagtcacacactggtctatctctctactctctgcaggaca  
gctgtcacactgatgcctcctgaagtaaggcagcccaaatgacagcatcactgacaagggttctgccccaaaagggt  
taacctggaatctaaccagcctctaacttcaacttctcatttgctaaaaacacagaggagaggggaacaaatttaataga  
ctccatgagaaagcaatgagacgagccagaaggcaggatattctacaggatgactgacctgttttttggtttttgtttt  
gttttcaaatgtcaagaagaaaaagaaagcaggctgggaacgggtgggtcacgcctgtaatccagcactttcacggatca  
caaggttaaggagttcaagaccagcctggccaacatagtgaacccagctctctacaaaagtacaaagattagccgggtgt  
gggtggcgggcagctgtagtcccagctagttgggaggttgaggcaggagaatcacttgaatccagaaggcggaggttgacg  
tgagccaaagatcgcccgccagcacttcagcctgggtgacacagcgagactctgtctcaagaaaaaaagaaagaag  
aaaaataaaacgaggagattattctagataggaagcagcaaaactacagctgattacacagctgatggtaaggttgctg  
cctgttggtttttgtttttgtttttgagacagggtctcactctgtccagccaggctggagtccagctcaotgcccacg  
taatctcccaggctcaagcaatcctcccacccagcctgccaagtagctgagaccacagccacgcaccaccatgcccgc  
taatttttgtagagacagagttgcccaggctggtctcaactccggagcttaagcgtccaccacctcagcctcccaa  
gtgctgggttacagacgtgagccgcccaggctggcaaataaatttttgcaataaaagttttactaatagaacactacac  
ttcattcacttacatggtgtgtgtttttgcaactaaaacaaagttaacttccaagtagttatgatagagaccatattggcctg  
caaagtctaaaccatttattattttggcccttcagagcaaaagtttgccaacccttggtcttagatcaacacacacttaaca  
gacaaaaaaaggataatcttcaggacaactgacctgttttttccaccatcagctgcaagaagaaaaataaaggagagg  
gtgtttcgaactaactatcaaagaagacattttgggggggtgcccggcgtgcccgtcacacctgtaatcccagcactttg  
ggaagccaaggcgggtagatcacctgaggtcaggagtttgggaccagcctggccaacatggcgaaaccccatctctacta  
aaaatacaaaaatcagctaggcatgggtggtgtgacactgttaatccagctactcaggaggctgaggcaggagaatcgott  
gaacctgggaggcggaggttgagtgagocaaagatcccgccactgcaactccagcctgggtgacagagcgagactctgtct  
caaaaaaaaaaaaaaaaaaaggccgggtgaggtggctcacgcctgtaatcacagcactgtgggagggtgagggcggcg  
atccactgaggtcgggagttcaagaccagcctggccaacatggagaaacccctgtctctactaaaaatacaaaaatagcc  
aggcgtgggtggcgcatgctgtaatccagctactcgggaggctgaggcaggagaatcgcttgaaaccaggaggcggaggt  
tgccgtgagccgagattgcccactgcaactctagcctgggcaaaaggagcaaaaactccgtctcaaaaaaaaaaaaaaa  
agacatttgggggatctggaaaatgtaaatggcctggatatgaggtaacactaagaaatgagtgatcatggaattgtat  
tatgtaaagaattatgtccttattttgtcagggaaacatactgaacaacttagagatgagtgacactggctcatcagtc  
agggacagaagaggaactacagggtggcctcatagtttagggctgggctgtggagggtgatctcaaggtccctctgg  
tgccaagctgctctgggagcagatgtggcctcacctcgctcccacttctcaatgtgggcaaaagtccaccaggcccaag  
ccctgectcatgggcagggtgttcagatttgggtaccaccaggagccccctttggccaatggaccaggccagcaacccc  
ctcttgagggtgagcttcaggcccccagcccagggtgcaggaagtggagctacacccattcacctctggccagggccact  
gtgggggtcagttgcctcagccctggaaagctcagcttgtcccagggaacttggtcttgagagcagcctgcccgaaccg  
ccccccgactacagctgcccccaaagctgtgaactcagagttatgacaaatggccaagcacaagagtggtgagggcctcc  
tatcctgcagaggctgctccagccacacctctgcagccaggaggcagtgacagccagtgctcctcaaggtcccccatct  
cccaggatgcttctgaaggctcaggttaaaactgcaaccagtggtggtgagcaaaaggctgcaggccctttcacaaaaa  
ctctgaactgggtccatgcagggcacccctccctgtcttctcccaactactgctgacccgtccccccgctgctttgc  
caaggtcatcccttcagccaggaacaatcccgctccctccagccacaccttcgacaggggtctcagccccacatttcttt  
tttaaggatgcacacccctgcaaatcccaggacaggactacagctctttttggaatccctttctcatgtctgggatcatt  
tgatatccctcttctcaccttacaagggtagtgactgtgtttttgtcaaacctgtgtccccggcactcagcgagcac  
cagacagggaggagctgtgtttgggtttatgtttgtgaaatgaatgaccacatcacatttgcttgaggggcctggccaagg  
ccgtaacttcagccctaaatgataggtcagggccagtgactgcaactgggggtctgaggaagcagaactccccaccca  
ctcttttgctcaccctgcccctggggggcacttccctctccctggcccccacaaagtggcctccactgggactcctca  
ggccttgctgcagtcagctgggttacctgtccatgcccgtgtttgagagggacatgccccatggccactgagggtaggaccc

FIG. 7

53/157

acatctcatgcagtgccagtgactcagtaaagggcaaacactgagggcctgtaacctctggatagtgacaacatagagggc  
aggaagcaagggacttcaggaacccaaaggaaactgggaaaacaaacctcctttctcaatggagaacctgctggtcgct  
gtcctcggggagctagactcttgtgcacacacaattcctcatcagaaacaggcaaaagagggcactgggcctccctgctg  
cagctgtgtccacagagaacagctaggccaccacacaaccccaagcctggccctgctccagcctcctcctgtgccaacc  
aaggcctcaccatggctcaatccacatggccccagaggcagcagtcctgggtcacaattcaggatctgcccctgtcagct  
gcatgggtgagttctttcacttctctgtcctgtttcctaatgtgtgaacacaggaataagcagtggtcctctccttccag  
ggtttgcgtgcaactgtgctgaagctctgtgacaatgcaccccccatctctgcaaaagcaaacccccaaaggcctggttt  
ctaaagcaacagcagtttcagcagcacctcagagaggcacttcaggccaatcctggaggagccaggagtgacccttaga  
gtgggccccagggacgtcagccttcttggaaaacagctcaagggtgagggggcctcctcctcctgctcccccttctc  
ccactcccaaagcagccaggtccctaggggagggctcagagaacagatgctgggagtttccagtccccctaacagaggggg  
tcacaaggaagatgtgcagaatgaacatcctgggaaactgggaaatgactaggagggaacatgggtgctccccccagca  
aaaaaattatacccttccccatggagatggaggtgcagcaagcttcaggccccagccccaggggtgcaggaagaggagcta  
caccattcactctggtcagcatgctcccaacgtctgagacctcatttctcattccttcttcagtccccattctcattg  
tggttcagaggtctttctctctgagccaggaaactccgcaaccttcaccctccctacctcttctcctcctacccagcctg  
gggctcagtcctggctcaagcactcagtcagcagaagcactgtgtagcctcccatataagctcaagcctgtgaaaagaa  
caccattgaggccttgagatggggccacactgaccgcctgactctcaggactggacacagcagagggccacacatactca  
gaacaaagcctggaaaggcaaggctggaggtcagtagttgtggcagcttcacatcaactcagctttaatgtgatttaatt  
tcttctcctccagtgggccaaagggtgcaaaagataagtaggtgttctctccttctaacagtgaggtgctgggggt  
gggggtgggggaatatggagaagggaacctcaccacccacaccttctgctcccccaacagtgctgcccctcctctgccc  
agcattctccccactttgcccctcagctagtgggtgcttagcctccagatagcatgccccacctaggccctgcccggggc  
tgtgatccagaggtcccaagaagcagagggccaggtggatccagggggtcagccaaggtgaggggtgggagcacacaggat  
tatctcccagggaacagggctgctgctcgtagctcaggatggatagaatgtggggggtatccagctacattttccctcc  
acaaaagaccagaatgggaggggatgggtgctgccccgacttcttcaactccccggagcagaaaaatgcccactcctc  
cacttccagtgccaagattcaagaagaaaggcaagcggagacttcccttctcagtcctctgcttactaatggaaacacg  
ggtccagaacctaaatccagtcctcctcctcctcacaacccgggagggaggtgcagcccaagcccccgagggcccaagg  
tccaggtgtaggacctttatcctctccggcagccatccctgtgggtgtggcaccocccgacacacccattcttgtcatc  
tcagggggaggggggaaatgtaatcggacatcccccccatccaatccatcctgagctgcgagggcgggcctctgtcctc  
ggagataatcctgtgcaactccccacttcaactcactcggctgacgcaggagtctcggagcccgcaactccagacatca  
ctgcccctccttccaggggtgctgaggtcggaggtcagagatgctactggtccaaggtcatgcagcagggcagcgga  
atgaacaggggtcgtggggaggagggggcgctgaccccatatagcccccgccctcactaccgcactcatgccccggacaa  
tcgcttcgcggaacacccccagctgccaccagttaacggtcactgcgccccgggaacctgacatcactctagttcaggt  
gctggggagtcttccagggcccgcccccaacagcggagccccccacccacagcctgcccacacggccgctgggggtccca  
ggcactgactgctcaggacagggccgggacggggagggggactcggcgaaacgggaaggaaacgggagggagggagcag  
aggggtggcctcacttgcgggggtgcaccccgggcggggaggggagggagggagggagggagggagggagggagggag  
gctcaacgcccagcagcaggaaggagcccccgccgcccggcgcctccgctcggcacgcccccgccctccgcccattgggt  
tgatctgggaggggtggggcgagggaagcctccggaccaatgagcggggtccaaagaacggccaactggcgagggcgccctc  
cgtcacgtgccaggggtcgccgagggcagcgccctgctagtccgcccctgcggggcgagctctcgcgaggaagacgggcagg  
cgccccactaggccaggggcccagaacccagccactcgaagagggagaaaggagggcctcggataggccccgccccgctcc  
ttcttcgcccgggggatagcgctctagccttgaaccttgccttaggacgcacctcccttgggccccttcgctctcgggag  
ggctgtcgggcgctctcgggggtgggtggagctcccgaaggtggccttctcctggggttccacgcccggcttcggcca  
tcgatacggggcggtgtgggtctcgttcaggagctgaggaacctccatcactcctgttccgaccccagggtttggacctc  
ttcccttccaccccatccctgtcttgaaagaagcaacccccgtgcggggcccgagacgctccgggtgctggcgggg  
cctgggagaagcatcagaacaaagaaggcacgcgggctgggggtgggagagcctgtgacgcgcccccggggacgcgagcc  
tctgctcccggtctccatggaggcgtcgccatggcagcgagatgcgctcgtcagcacgcgggggtgggatgtgggag  
cctgcaatgagccgagggagcagagggcgtggcctccgggtctgcgggggtcttgcgggtgctctccgcccgcgggttcg  
cgaacaccccacctatacttcgcccgtggggacggattccccaaagtgcctcagtaagcgtccggagcacgcagcgcc  
ttgcttccaaacggaaactagagagacggcctggggcgccgaaggccagcctccctcagcagggccgggggtcgtgctta  
aaggagcctcaagtctgccacctgtggcccataacctgtctgctgatctccagtctgcacactgttggcaaatataatc  
ttctgagctcttggtttctatcgcgctcctctcctgctccaaagcctctgggactgctccagttagcgttcacaaact  
tcagcagcactttgggtgactcatgtgcccctcgcggttgagagacagcgttacattgagagcttttccattctgatct  
cagcccactcactcctgcccactctaccactccacccacactaccttagacctgtcttgaaagcttaaggttggcag  
acggtggggcccaagtgggccgttaggtgtattttgttagctcacatgggggtttttaaagaagtccttccccctcctg  
cctcctccgctctttctggaatcaggaatctgggttctggtggcttttgaaatatcagaagatctgacaacgtggggtcac

FIG. 7

54/157

[illegible]

FIG. 7

55/157

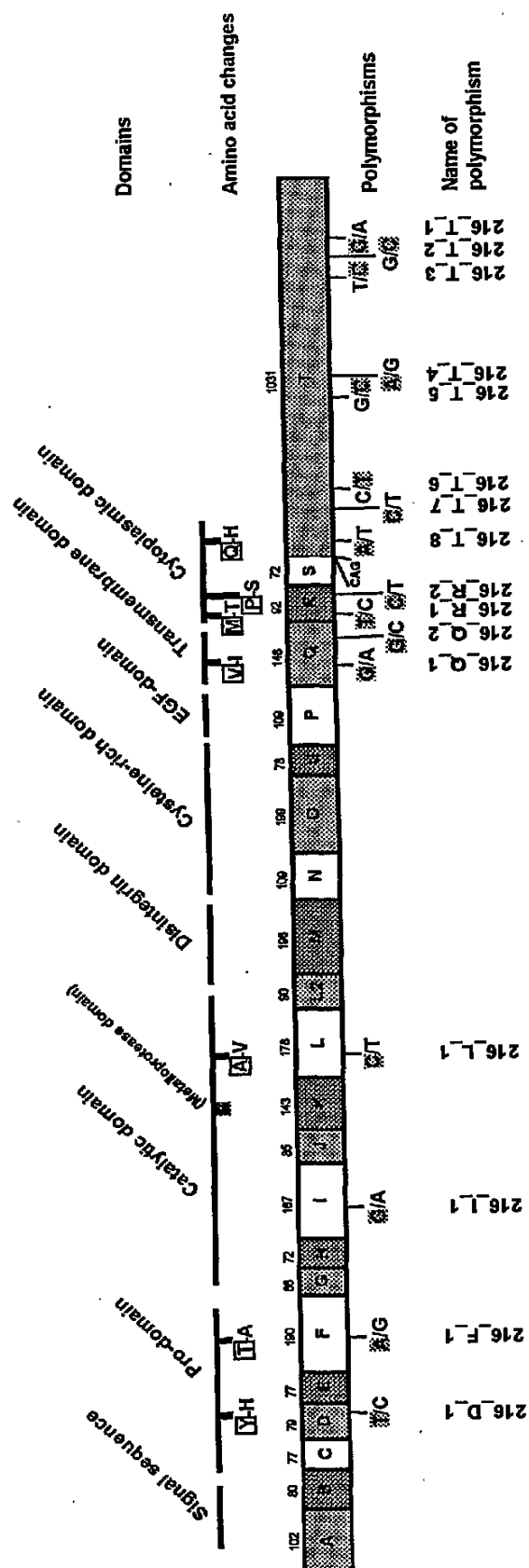
ccacgagggtggagggtgcagtttaactgaattgtgccactacactccagccagggtgacaaaacaagactctctctca  
aaaaaagagtagaagcaaatgtgcatgtttgtttatagtcacataaaagtcatccggaaggatacataaaaactaaagat  
agagggtccctattagagatgtgctgggaataagaagacttttcaactttttatgtgtttttttaaatcatatatatgtt  
tctctattcaaaacatttaagtaacaaagtttaaatacccttgcatgccttctctcctgtagggttgtgaacactggag  
ggacaaaatcatgttttatactcacttttgtattcctgggtatttagtactataggtactcaataaatagaggaaagaa  
cgaatgaatgaaccagaagatagcagatttagtgtcggttagagagaggaggggcccaggccagctggctcacacatgt  
atccccagatcttttggagggtcagggaagaggctcacttgaacctaggagtggagaccagcctgggtacaaaagcgaa  
atcctgtctctccaaaaaataaaaaaataaaaaaataaaaaaagcccaggagcagtggtcacacctgtaatcccagca  
ctttgggaggccgagggtgggtggatcacctgaggttagcagttcgagaccaactaggccaacatggtgaaacacctct  
ctactgaagatacaaaaattagtgtgggcgtagtggcagggtgcctgtaattccagcaactcaggaggtcagggcaggagaa  
tagcttgaacctgggaggcagaggttgcatgagccgagatcatgcgcctgcactccagcctggatgacaggtgagaca  
tctcataaaaaataaaaattagctggatattggtggtatgcacctgtcctcccactattctagaggctgaggtgtaaaga  
ttgcttgaacttgggaggctgggttgcatgtagccaagatttgtccactgcactccagcctgggcgacagagtaaaacc  
ctgtctctaaaaaagaaggaggttagggctgggtgcggtggctcacggctgtaattcccagcactttaggaggccgagggt  
ggtggatcacctgaggtcaggagtttgagaccagcctgaccaatatggtggaacccctctctactaaaaatacaaaaag  
tagccgggtgtggtggcgtgtgcctgtaattcccagctacttgggaggctgagacagaggaatcacttgaaccaggaaagt  
ggagagtgcatgagccgagatcacaccactgcactccagcctgggtgacagagcaataactgtctcaaaaacaaaaa  
aagaagaaagaaaaatttttaaaaataaattaaaaataaggaggttataatcttatctaatgtaatcacaaaagttg  
acatcctatcacactttgctgtattctacttatcagaagcaagtactaagtcagcctacattcaaggaaagggtacaca  
tgggggtgacgtccaggaggcagggttactgggagccacgttaggagctgtctactacaaggaggaaaaagcaatgaaa  
ggatgaaatctaggatggtggttatctctgagtggaagcaaggggagagatgcaagaagcacatggccgggtgaaagtt  
tttgtaataattccattcttggggcagggtggtggatttatagggttctcataagtgagttgatttgacaaaatggggaga  
gagtgctcatgaaccaaggtatatatttagtccaattcagtaatatttaaaattatcaataataatgaaataaaaaattgc  
caactactcctgacctgttctgagcccatccatgcactccacctggctcttagtttaacactggctccacctgacctc  
cacttgcocccaaactaaccttagccttccacacacactgaccttgctatagcaccacgatctctccatagctctgcaat  
aggcccccggagttctattccctgaaactttatccatgtaacctcacctgtaggctgctatatgtaaagatagtaggaa  
tggctgatggagtggcccgaggaccaggagacagggcaggctaacagttctgccttagcagtagcactgtgggcccagaa  
gacaggttgatggctggtatcacgacctcacctattggcctctacatggcaaggctgggcccaggctcctgatggttct  
tctcttttaggctctctgctcctggtatccagatatagctaaagtcgtctctcgagaagtgccactgtccaagcaccag  
gcttgaactcccttgttggttaataagactctcaactctctccttctctgtcttttgatgaagtaagctaccatactgga  
gatggctatgtgtcaaggaaatgagagtggtctattggccaacaatgagtaaaacactgaagctttcagtttgacagocca  
ctaggaagtaaaattcagtcagcaaccaagtaagcctggaagcctggatccttctcagttgtgecttcagatgagactgc  
agttctattggcccttctgcttgcagccttgtaggaacattgaagtgaggaccagccagcctggcctggactcttga  
cccacagaaaaatgtgagatataataaagtaatttgttgttctaaagctgtacattgggtgataatttgtcacagtaag  
agataatgtatacaggttctcttctattcttttgtattacaaaagagtggttctgtgggcccaggcatgggtggctcacgc  
tgtaatcccagcacttttggaggccaaggcagggtggatcacttgagggtcaggagttcgaggccagcctagccaacatggt  
gaaacctgtctctacaaaaatataaaaaattagccagatgtggtggcatgcacctgtaatctcacatacttgagaggct  
caggcaggagaatcgcttgaactcaggagggtggaggttgcatgagccgagatcgtgcaatttgcaactcagcctggacga  
cagagcaagactccatttaaaaaaaacaaaacaaaacaaaagagtggtgtctatgaacctgttgttaacctgtctccca  
gaacacagaacacaagagccaggtttcccttgggtatgtgctgggttatgctcaggaattggaatttgttagctcagg  
atgctttgttttgttttgagacagagtgctccttctgtgctcaggtggagtgcagttggtgcgatcttggctcactgcaac  
ctctgccccccagcttcaagcaattctcctgcctcagcttcccaagtagctgggactacaggtgcccgcaccacgcctg  
gctaattgttctgtatttttaataagagatgggttttaccattggccaggctggcttgaactcctggcctcaagtgaccc  
acctgccttggcctcccaaagtgggtgggtatcacaggcctgagccaccatgccagctggatctcaggatttacaatgtt  
ccacattactggtagtgcagaaaaccattaaagttagggtcatgccatggctgtagtgttggccagggttgggagtaaa  
gctcggcctcaatccatggaagtgggtattgtgtgggcagggttggagtgcacagaggccaggttaggggtcagtc  
cttacgtagggttgagggtgagtgcttggccacatttgggaacaatttgaagtcgggttgggttgggttgggacaagg  
ttggagaaactgtgagggtgggttgggggtcaatctaaggccagggttagaagttaggctgtcatcctgtgcaaaacc  
tgacctcctgggtgatcatcaacctaccaagctgtggccgcacaggaaccagccaccacaagatgagatccactctggg  
tccagaaagctcttctcactcaaggccgggggtggggagatggaggacaagtgaagaagaaactggctccctcagatgcaa  
aaagaaaggcaggaacaggaaatagaacacaacactgactttaatggggcagcctgagccgtaaccttcaaaccttcc  
gctcagtaacccctgcgcgaggaggaggggcgactacgggcacatcgtcgatgcctcctcactcctcctggcgatcc  
cacgctcaccctctcctctccagctgtggcggtctccgtcgccagggtgggtcccaggtaagctcccaggaggggcagg

**FIG. 7**

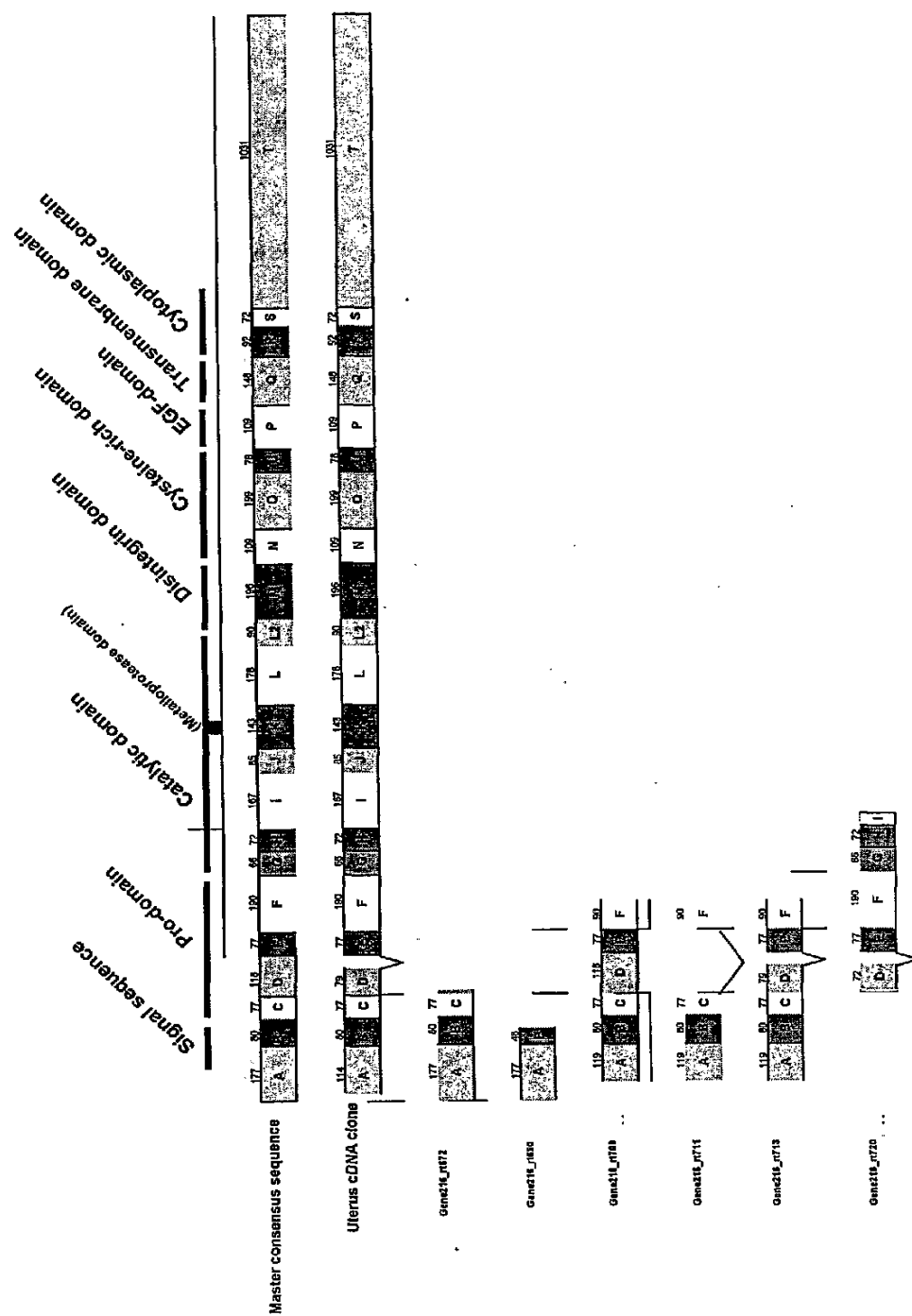
56/157

ggagccctcagcggccgggacagattccccatggctgagggcatggggaactagcctggatggagacgcgcgcgtcc  
tcggagctgggcggggacctgatttctgggggtgtggacagagccaaaggaccgcccccaaggccagcgcgcgggagatg  
caaggccgggcccagggtgttggacactgcttgggggacaggctaggtctctgcacgtggcttctgggctctggaaagc  
ggtccattctcctgacccggatctccggagtggttaggaggcggtcagtcggggcctgcgctcctagagttcctgtccc  
atctcctcccacgctcaccatcccaaggaaggaggcactcgggcccagcaggctcgtgagagcagcgggctccgccc  
tcccaatggtctatccatcggtgggtgggtcgggcgcggcgtcggggctctggcgggtaccggggtccccgcgcggcc  
cgggcccgcgtcaccgctgcttacgctccgctgctggagccgcgggagaggtcttcgatccgcacattcttgagctg  
gagcaggtgctccaggcgcttaccttcatctgcaggacctaaagccgcacccgcagggcgtcaagcctggcgggtctgctc  
cctcctgcccggcctctgctggccgcgagccccaccgcagccttcatagaccctgacttgcacggtctcttgagagg  
ccaacagctcctgctccttttcagcgatctggaggacgaagctggggctgcgctgcaacgcgggtatacgtggaggc  
cgaggcgccggcgggcgggtcccagctggggaggagcagcttagcggaggcttggacctcgggcccctacactcacctt  
ccaggggccttagaccaggccttcccttctccccctctagcccttctcgtcctccgctctcaggctcagcgtaccgtgc  
ccccaccccagctgtgtgcagagatcaccagcctgggtgaagcttggaactccagcttctttacctgagctgacccct  
cccggggggtccgggacaccttccctcgggcttctgggatacacctgagacaaggcaagtgagggaatgactgtgctg  
gagttctcattctcctgcttccccctccctacttccactctcacaagtgcaggaagaaggacagaaggaaagctccaacc  
gggaggggagaaatggaaaaggccaccttaccacatccatgtagccactgccatcctggggagccagctcctgaaagaac  
ccgggggtccgctgttaagcccggtcagtcaccccaaggagcccatgtcccttccacgggccccagggcaggctc  
tattagccaagcaccctccccacacttatcaccaggctggccttctgtggtctggggagcgcggccccaccagccc  
ccagagcagtgaggagaagctggagcgggtggcacctgttaaggagccgcgcctgcttctgcccctctgctctcctc  
caggcgctgctcagcgggtgagcaccagctccaccagcctggggcgactgcgcgatcttgccatcacgtcatccg  
gtactgaaaagttcagcctctcagtaacctcctgaggggttagacattgggtagcagattgggtcctggggaatgaag  
cgaggggtgaggggagagggaagggggcatggtggttgcgttctgagtgtagcagccctgggggagtgaaagcgatcgc  
atctgagcttcttgtgtgggtgttccgaggttccctgctccttctcctaatagaagagttcttgattttgtaatcggac  
gaattctgtctgcgttgaccctccccaccctgccacaaaacagctctcagcctgggcaaccaagaccataggagatgatg  
gggctgcacaggtattctgctacctgttcagatgacccagttgctgagcttctgctgggagagagttggcggggacataa  
ttgtgcatctccaccatcttggggaagttaaacttgatgacctctgcaacaaggactgaggggagaggggagcagacatgg  
aaagtgggggcatgggttaagggaaggaggaggtgagagaaggtgggggtgggagaatcaggcatggcagagtggtggaacc  
caatggaggggtggagatgcggaagggaaggataaggaggagacaaagacatgaagattagtgataggttagaggaga  
ctaagttccagcatacctcagaggagctgcatccctccccactaattaggaatgggaatactgagttctgtgctc  
cagtagagaatgagtcagccagtgaggggctcacacctgtaatcccaacactttgggaggtgaggcgggtagatcac  
tgaggccaggagttcgagaccagcctggccaacatggtgaaaccccatctctactaaaaatacaaaagttagccaggtgt  
ggtggtgcatgtaaacccagctacttgggtggtgagggcagaagaatcgcttgaactcaggaggtggaggttgagtgag  
ctgagatcatgccaactgactccagcctgggcgaagagtgagactctgtctcaaaaaaaaaaaaaaaaaaaaaagaga  
atgagttccaaagagaagcttccaggagccagggaattcacatggatgcatgatgcacacacacacacacacacaca  
cacacacacactccaaacccaaagttccaggagagtaaaagactcacgttaagcagagacacatgtgataagatcccccta  
cacatagatacatagggcagaatgtgtagacacatagaacaaacacatgctagacacacataggcacaacatgtagaacac  
acatagacagaagcacaaaaagccacatatgcacacagcctccccaaatatacagatacagagatgcatagaagtacttg  
acatgaaaagatgacagataagttttctgtcatttgagatccaggagacattcctagaacacaaaacaccttcacaatc  
acaaagaaacacatgcacacaagaccccccaataacaagactaaacatgtcaacacatatgaacacacacatgctcaga  
ctcccagaagacatagaacacatagcaatgagtgagcacatacagaaatgcataaatacatatgcagacacatatcttaa  
tcattaatgtccctggaggaggttagaggcacccccccaacatgcacacaggatagaacatgtagacatacatatatgg  
acacacatcgttcacacacagatgcaaatagaaaacaagcacaacacatctaaaatttgaaagtcocaggaggcagattcct  
ccccactcccacaagcacacaaaacacaaagagttgcacagagaccogttacacacgaacatacaggacagaacatgcag  
acacacaagggcacacacacctccatcgctaaagtcccgggagaggtttcgcttgggcggggacagagggtgtgtctacc  
cacaggtacagctgggtgcagcgctcctcgctccagctgctcgccattggcgtcctcacggcctggccgccccagcggtg  
ccgggtcccgcggccagcctcacgaacctcaggcgcttcccttctcgccactgcagaagccataaactgcctctgctgc  
ctaattccagagactcacgtcccagctgggagcggccatattgtttctgaaacctggaaaccttcccgactctcctgg  
gtacaaagacttctgggagttgtagttttatccccctccccgctacctcctctcaggccccagctttgaaaaggcggttag  
ctgcccctgtctcttcccaagaaagcatcacttctgtgccacgcacacctagtgaccagccactcatccattttgggac  
cagccaaagccaaccaggaccccagaattccgcggtccttctatagtgacactaaatgtcgacggccaggc

FIG. 7



**FIG. 8**



**FIG. 9**

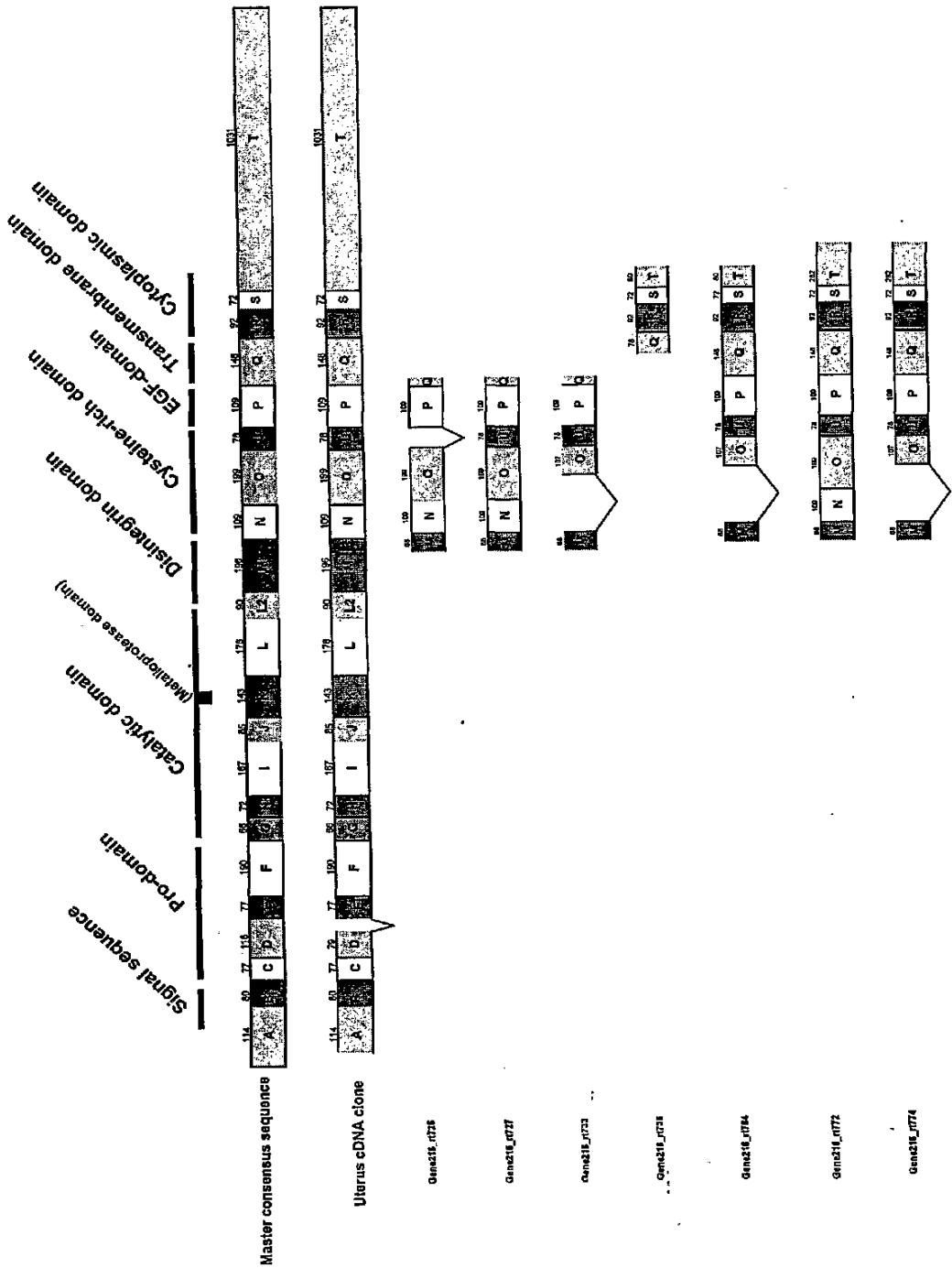
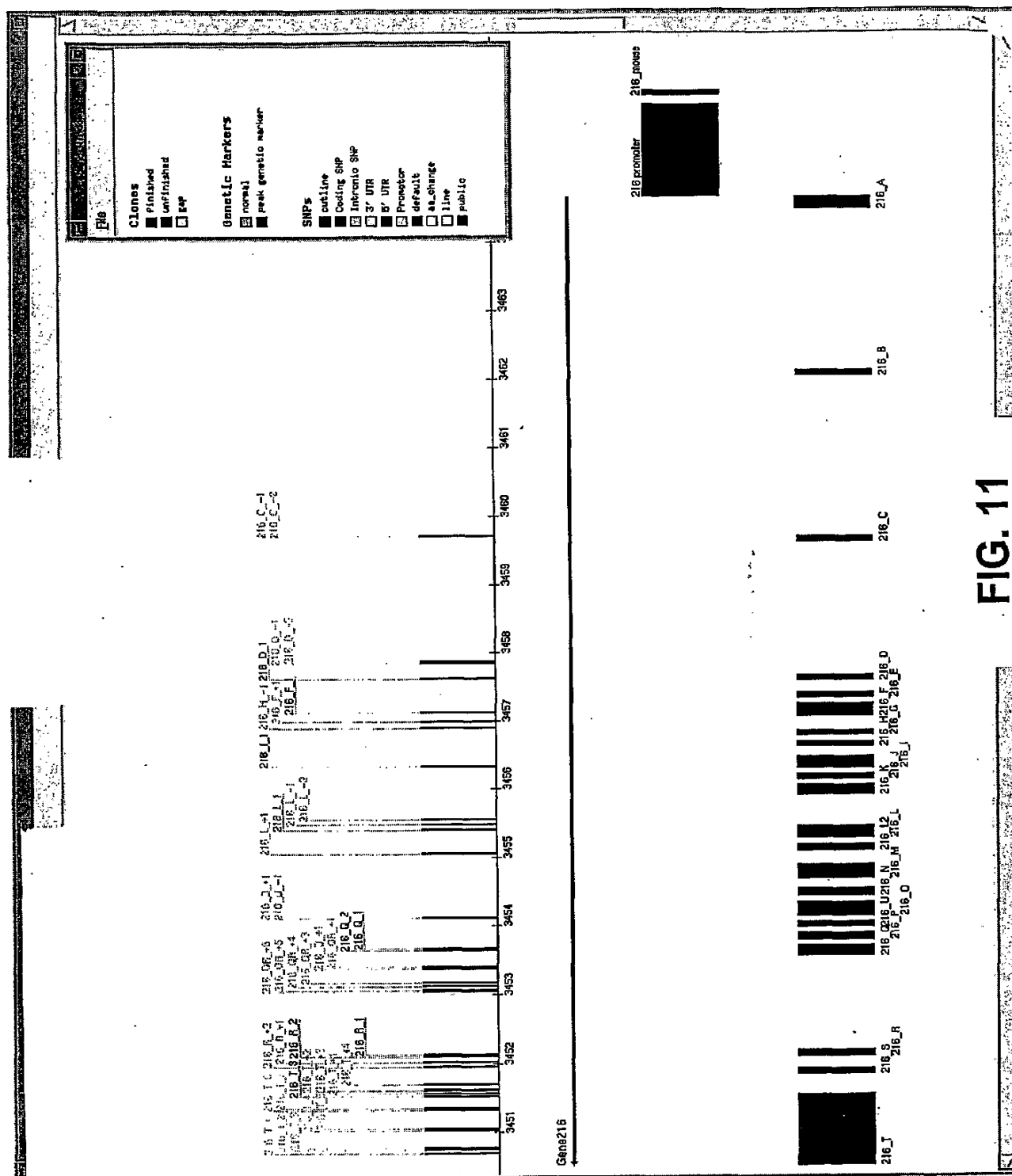


FIG. 10





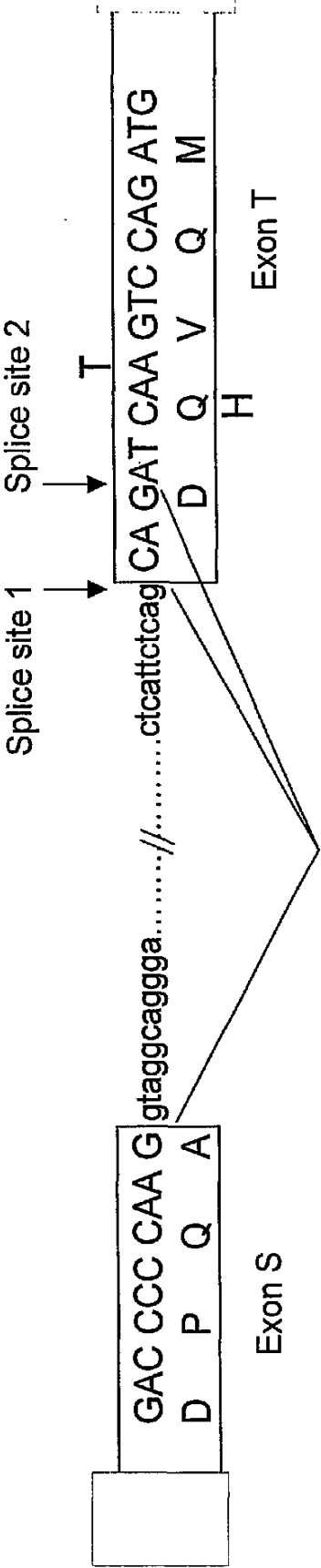


FIG. 12

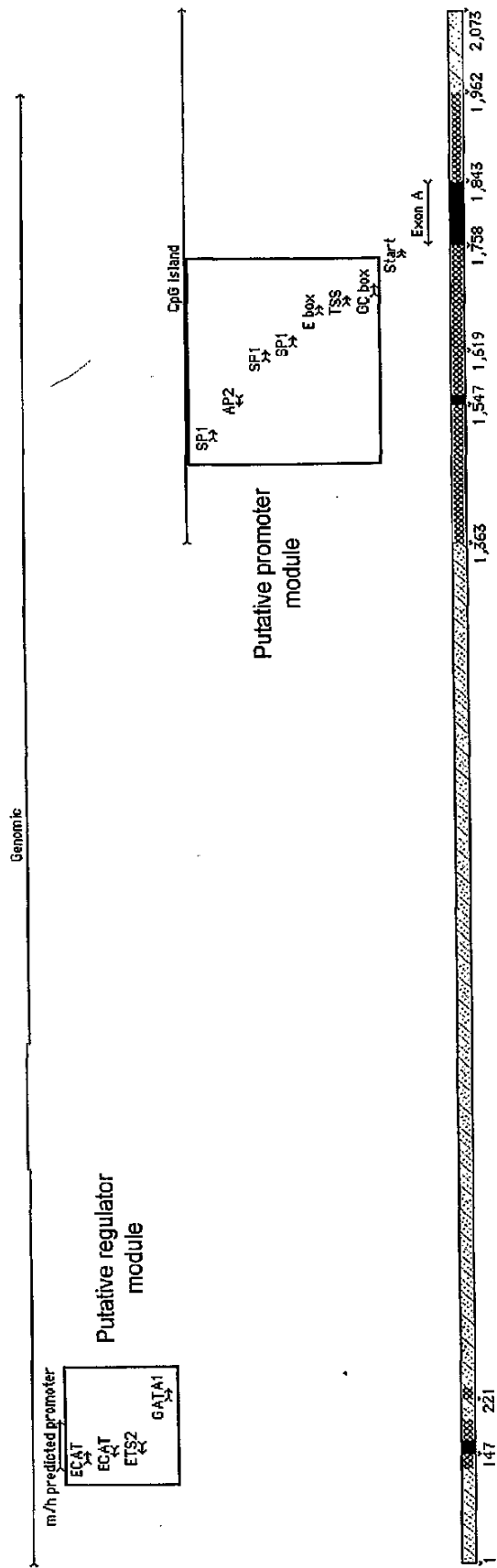


FIG. 13

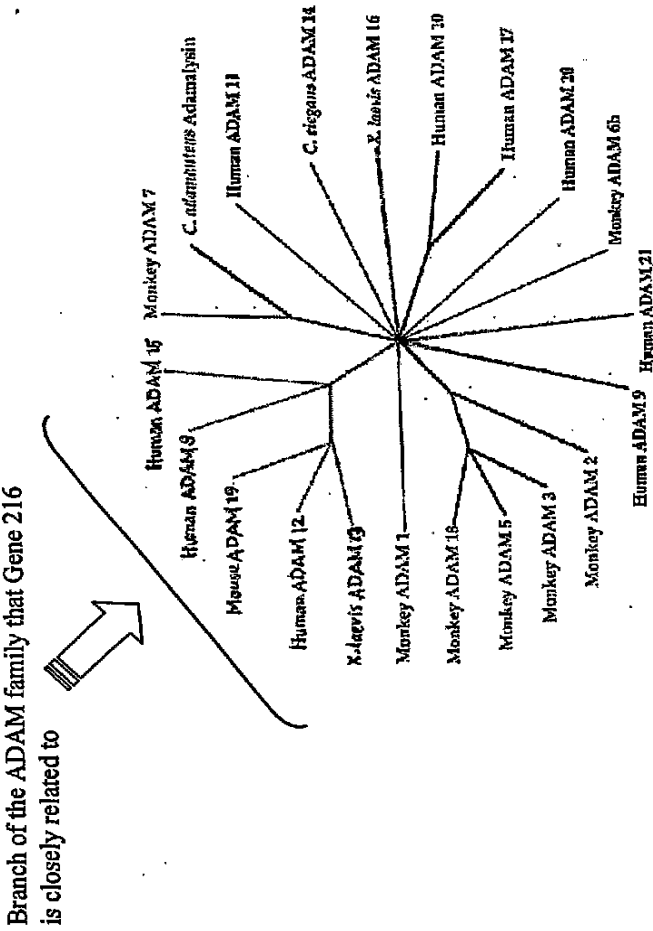


FIG. 14

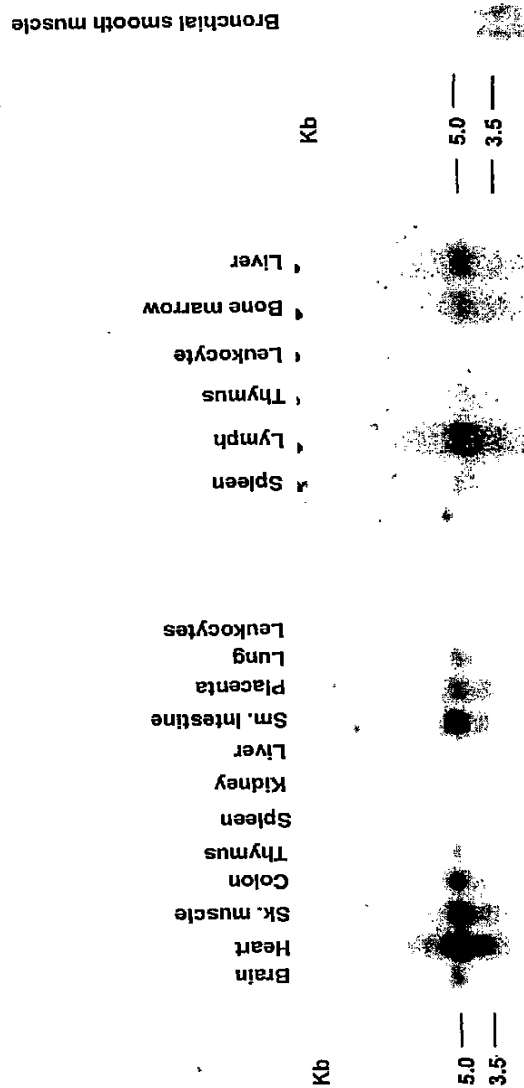


FIG. 15

	1	2	3	4	5	6	7	8
<b>A</b>	Whole brain	Amygdala	Caudate nucleus	Cerebellum	Cerebral cortex	Frontal lobe	Hippocampus	Medulla oblongata
<b>B</b>	Occipital lobe	Putamen	Substantia nigra	Temporal lobe	Thalamus	Nucleus accumbens	Spinal cord	
<b>C</b>	Heart	Aorta	Skeletal muscle	Colon	Bladder	Uterus	Prostate	Stomach
<b>D</b>	Testis	Ovary	Pancreas	Pituitary gland	Adrenal gland	Thyroid gland	Salivary gland	Mammary gland
<b>E</b>	Kidney	Liver	Small intestine	Spleen	Thymus	Peripheral leukocyte	Lymph node	Bone marrow
<b>F</b>	Appendix	Lung	Trachea	Placenta				
<b>G</b>	Fetal brain	Fetal heart	Fetal kidney	Fetal liver	Fetal spleen	Fetal thymus	Fetal lung	
<b>H</b>	Yeast total RNA 100ng	Yeast tRNA 100ng	E.coli rRNA 100ng	E.coli DNA 100ng	Poly r(A) 100ng	Human Cot1 DNA 100ng	Human DNA 100ng	Human DNA 500ng

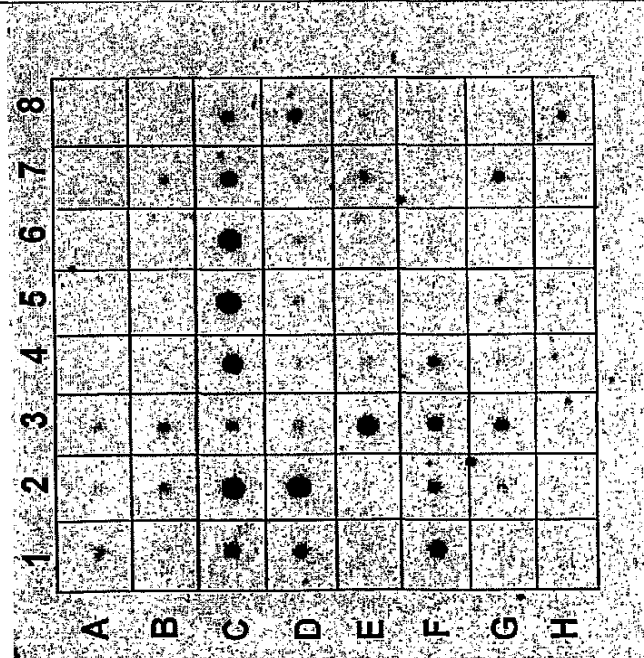


FIG. 16

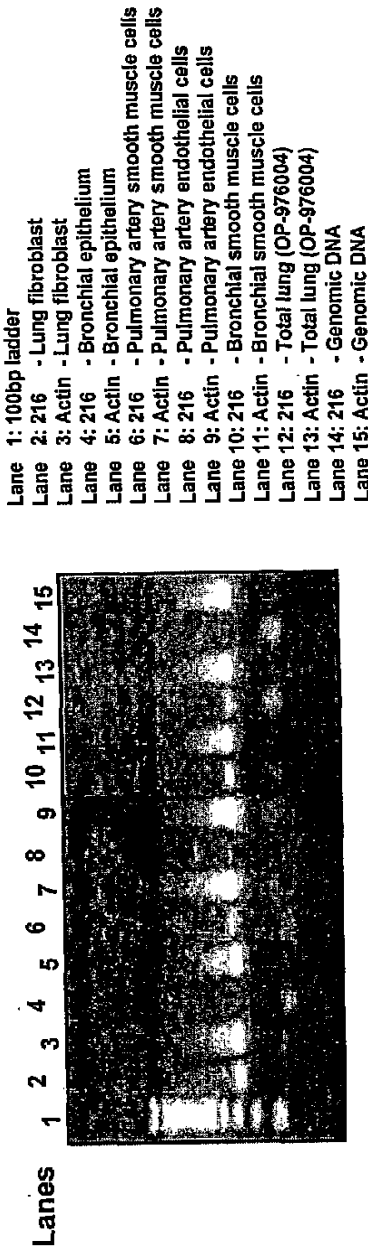


FIG. 17

67/157

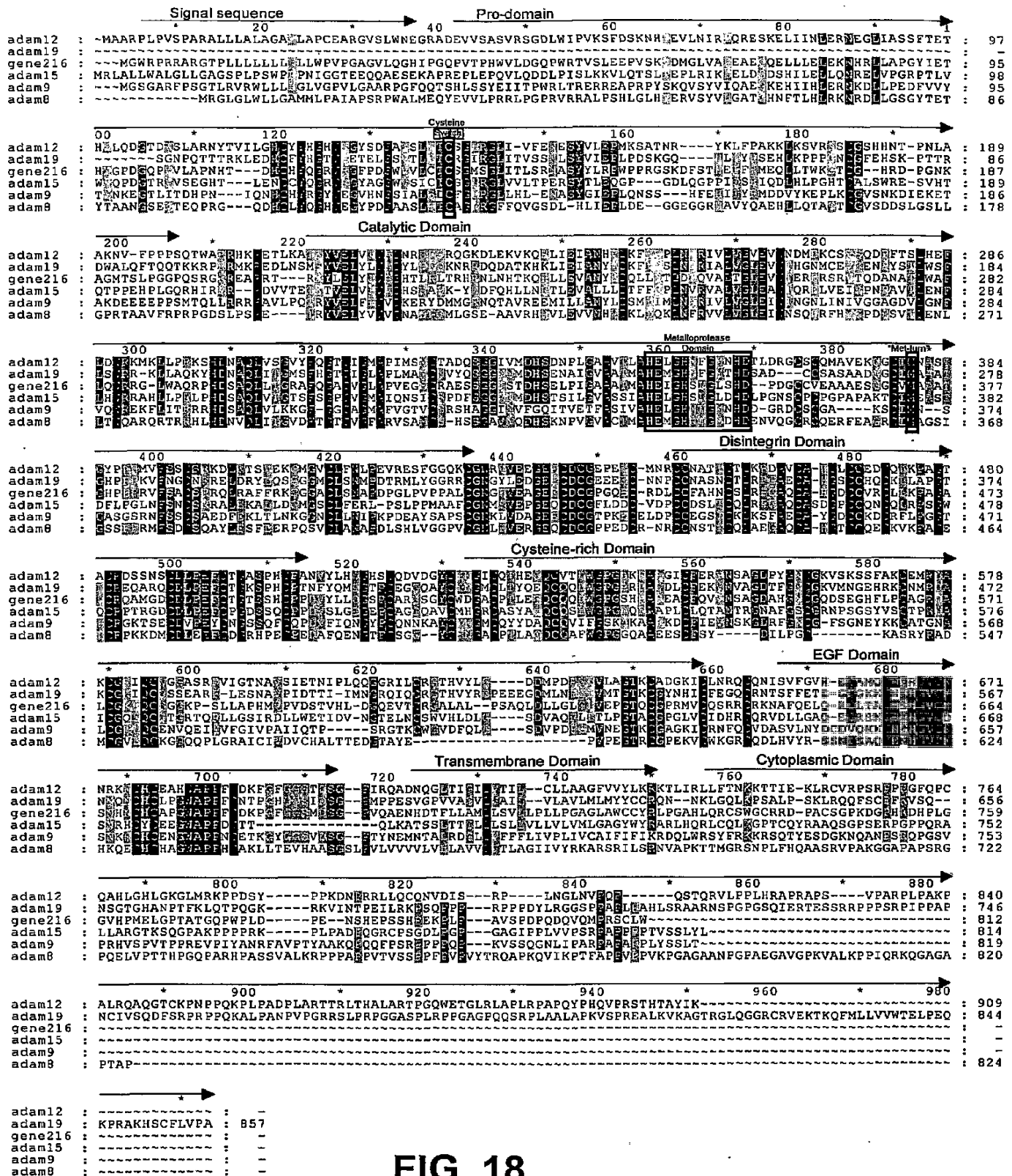


FIG. 18



68/157

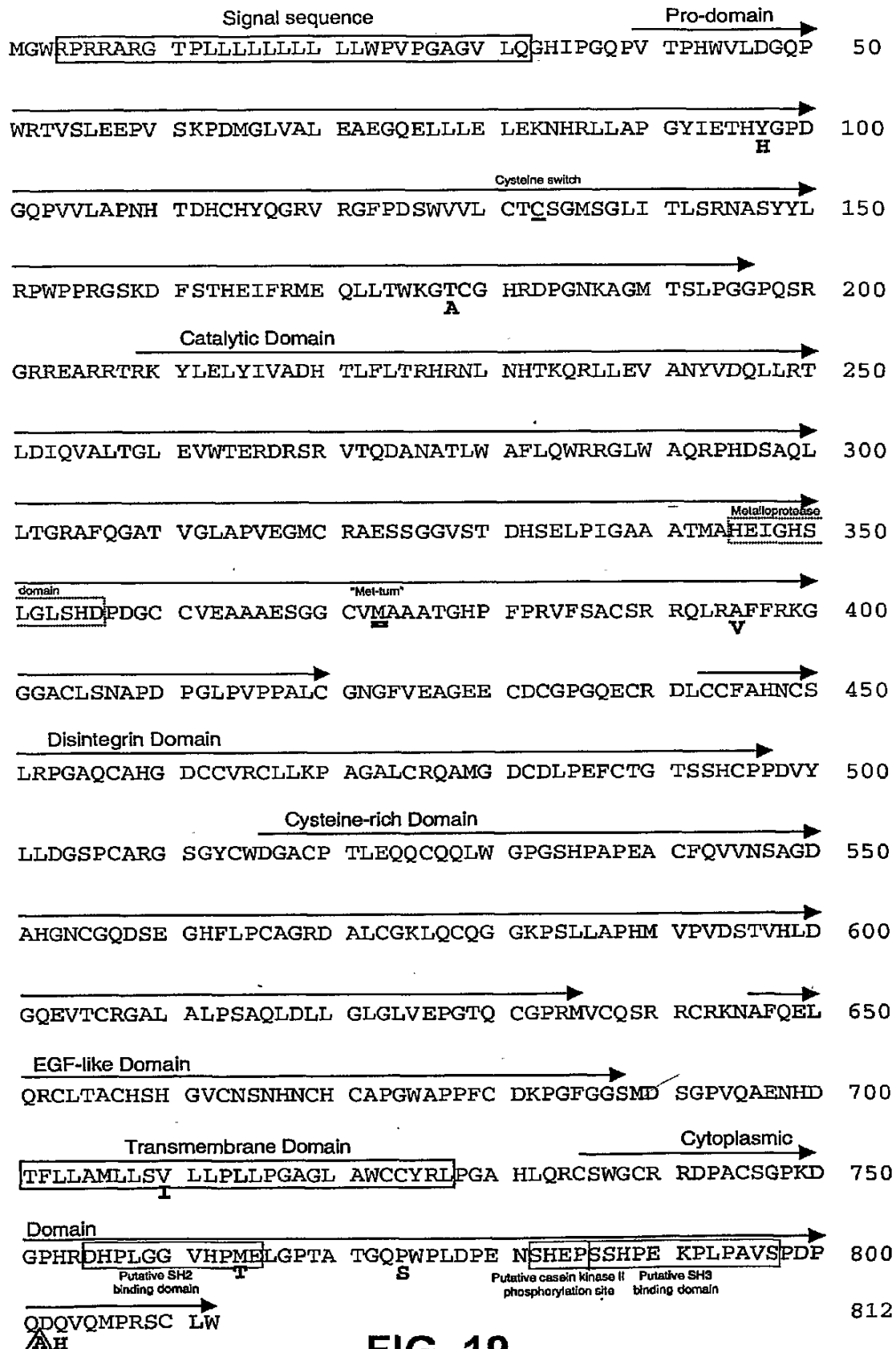


FIG. 19

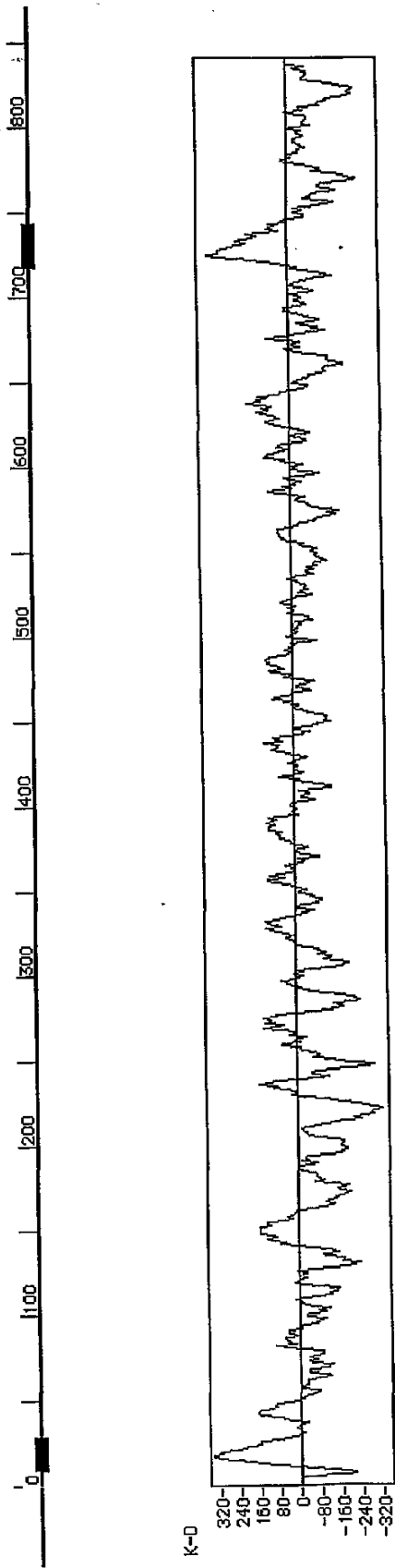


FIG. 20

70/157

ggagaatagtgggggtgcagtaggaggtagaagggttggaatgaattggaaaaggagtgga  
ctgtgtagccaaggctgtgccttgctcattctaggccaatttcttctcttctgtctctcc  
ctccctccctccctccctccctccctccctccctccctccctccctccctccctccctcc  
ctcctagctatccttccctccctccctttacttctatcttcccttcttcttcttccctcc  
tcccttctcttccccattccccctagaccagggtcccatgtagcccaaactgacttcaatc  
tcattacagatgaccttgaacttcagatccctcagtcctccatttctttttgtgcctgaata  
tcgtattttctctgggagcctgtccagtatcttccccaataaatttcattccaggatac  
aactggagtgcaattttcttcttgcaacccaaatgtctaggctagaatgtggagaattcaa  
atattaaacgctagtgctcagacacccatctgtgggcatgtggccaggaggagcataaatcg  
cacagtgccttgcttcatactgtctgctgaagtcacccgggccccttccctgctcctcatag  
ctcctcacatgataactcacggggtgggctgctcgtccacagggactggcacaggcttgg  
cacatggcagggtggcacatgggtcttctgtacatgggtgaatagaagcacagaaacccagct  
ctaccagaggcaagacaggaggagagaacaaacccaaagatggcaagatgttccaggatgacc  
gagagcaaaagccacacctaaggggtgggtggctgctgaggggttggggacatgtgacttg  
aatctttgctctcaagtacttccaagagggtggtggagactacttgggggtggggctcat  
tgggcccattctcagtagtctaaagcccctcagaaataattacttagctgtaattcaag  
cacaatatatacacatggtaagtgtattagcaaaatggctttctatgtaatatataaag  
ataaaaaaacaacaaacaaaaaacaacccagaaataccagcacttcagagtagcccccac  
tcccaacccctgcccagtgtagctgctaacttcactaatcaagttaatttaaaaaatcaaa  
tataataccacagatttttatttttccctgtgtttctgctttatacaaatggatatcatattt  
taaacatgagaacattgtgagatgactagccagtgctgtctgtctgtctctctcagata  
tttattttattttattttcatgtatgttaggtacactgtcactgtcttcagacacaccagaaga  
ggacattggatcccattacagatgggtgtgagccaccatgtgggtgctgagaattgaact  
caggacctctggaagagtagtcagtgctcttaaccactgggccatctttccagtcctgac  
cagtgctctcttaatgggtgaaatattcttcccttggccctagcattttatccttgggcatt  
ttctcaggaaaaatgaaggcgctgtcacccctgagatttgttactaggatgttcataaca  
gtgctacccatcagccaggaactgaaaacaataaatgctcggcagcaggggaatgaccga  
gtagacagagctacgcccacacagctatgagagctgtgcctacagctgagtggttctc  
agacggctcagtgaggagggtggcagcgatggctgcctgtggcagttacatgtcctaccg  
ctctattaaaggagctaacaaggaaccccttaaggaaggaaacattttttgggtcacg  
gttagtaggcacagctctcagagtggttggaaagatggcagcaagtgtctctagctgaggg  
ggcaggagtgaggaggccctgggtcactttgtgtcggcgaggaagcagacagagaggtggg  
gtgccgtttgcttcttactttcctcttccctaccagtcaggatgacagaacatcaggta  
ttgccaccacatttggagtggttcttctggtctctgtttaaactctctggaaacaccct  
cactgacctgatgtgcccgaacaagggtgattctcagcccaagcaagtcacagcgaagac  
caaggacagggtatagtttaaggaatgggatagaagtcaatacacagccactggtgggcaaa  
agggatgggcgaggagcacaatggagattcgagagccagggaagtgagcaggttcaacttg  
tgaaaatctcacattttctgaatctatatatttttgaaagaaaacttaatttaaaaaaatctg  
ccacttttccctacgccttgtgctgtgtatcgacatgggaggataccttgaggccaggaga  
ctgagatcagcctgagcaacacaataaaacctcatttcttataaaaacagaaacaaattggg  
tgtggaggtgcacgcctttaatcccagcaccaggaggagcagagccaatcacccttgtaac  
catctcaaatcagacccaggggctgtctagcggcagaagagtcctatggaacactgtgac  
catagtcatacaattggaaactacttatcgacaagaataaatggccataggtaatataggca  
ataggttgggtaaaactcacaattgggaaaatacaaccatactccagacaaggggccaa  
tcgccttttaacatagaggtccttttaactgtttaaagaaggaccaaccactcactagaa  
aatgagctacacatgaatatgttactatatatttaaaaaaaaaaaaaaacctgtgtctaaaa

FIG. 21

71/157

cagcagcttatgtaagaaattggtagtagtggccttctgtggggaaggagtacccaaagg  
atggagactaaaggaagactctoctaagcattttataccttctcccttggttgcttgg  
gtagtgggtgggtggttgtttttgattctttttgtttgggtgggttttggttttttcaaga  
cagtgtttctttgtggagccctgggtgtcctggaactcactcaggctgccttgaactca  
gagattctcctgctctgggcttaaagggtgtacctcaccactgtcaggcttgggttgg  
tttgagacggggtcttgctgtgtagctccagctgcctctgaagtgtagaattacagaca  
tgaactacagatatctcccagctctcctctcattctccctccctgcccccttgtatacac  
attcacctgtgtgtgcatgtgtgactcactctgagttcttccagtgttactctccatcag  
attttttgagataggggtctcttactgaacctggacctcaaacatttgccctgtctctgtcc  
agtccaacacttggcttacaagtttgtgcaatgggtctggcttctctgtattgctggaga  
cttgaactcaagaaccaacgcaggtcctttacttactaagccttctccttgatgacaccc  
ctcacccactgccacacaccaggtctccaggtgatctcacactttttccatataggac  
tttgaacttgtcttttgccctctctgctgggatacaggtgtgtgctgtacacccaattta  
tgggatgctggggatggaactcagggccccaatcttgatcaactgaggtgcgctcgctc  
tctcgctctcattccc  
ccctccttcccttccctctctcttccatgagacttctttctatattgcacagggtagcta  
tcttgagtgtcctgtctcgatccaattgcctgacctatcatactgggtttctgtttttgg  
ttttgggttttttcgagactgggtttctctgtgtagccctggctgtcctggaactcacttt  
gtagaccaggctggcctcgaactcagaaatccgcctgcctctgcctcccgagtgtggga  
ttaaaggcgtgcaccaccaacgcctttaatccagcatttgatgttcgaatcgataaatg  
ttttttcttctgcggttggaaacttttacattccatatctcagaacgaaattattaatca  
tcttattttcaaagttaaatctcttgctatccgaaatcataccaagtgtacgaagtgtct  
tttcaacacgaaagcaatgtccttaaaatttgggaaaggggtccctcatgcctgggggtct  
gtgggtatagaacaattctgtcctggccatcgctgttctctgcctgtgcccaccactgc  
attgcctttaccaggaaccacagacttcagcaacgcggccaccacaagacccttaggtac  
ccagatgcatctcagctgcagtgggacagcttactgctgtgggacactaggtgcaggta  
gacagatcctgggtagaaaggagaatccagatgtcactagggagactgagaccagagatg  
tgtgctaggcagacagagcagctgtgggatcacttagctagtgcctctattcagcacca  
gagaggttctctaagaaccttcccgactttccagcccatcagccatcccttggccagcg  
tggataggattgttgcagataacgccaggagggtcatcctggctcccgagcgcgacat  
cctgttttgcgggccagggttagggccacactcctgcgctcctgtgaggtggcgcgaggcc  
cgctgagggcgggaggaggcgggccctctagccggggtggagccaacagcacagcaaga  
gtgagaggggaaaggcactccagccatgggctcgaggtgcgggagaccgggggggtctcc  
gggtgctgctattgctgcgctgttgcctcgtgtcgcgtgcggagcgctcggtgtt  
tccaggtgaggatacaggcgggggggggggggggggggcggtctagggcagggtgaag  
tggtgggtggtcatgggttccagtgcactacagaacatagagttcccgaccctcggctac  
ctggccaggacaggatgtgggaagggtctgtgctcctggtaagggtggccctggagcaga  
cttggccacacttacttagcaaaagcaaataggcaacactacccttctcctcagtgtgccc  
agctctgtccggaagcttctgcacttccaaactgccaggccagagtgtctccctgctcct  
ggaaggacctggggtccttggagtccctgggtctgccacaggcctagctgtctcattcggc  
tcgtcctcttctcatgctgggtctgggcagtggtatgggtgtttatacagttgccttatg  
tggcagttcttcaaacaggggtgtctgcctagccgtgcctggggctaccactttctcttc  
ctactgtaacctctccttactccgggacgacagggctcttagcatctccaagcctgtagg  
gtgaatctccaaactccacatagcaagcactaggctggccagaaggcagaaggaccaccg  
gggaggaggcagtaagggtgagggtggatattgggcagctctggaaaccaccccatatc  
tgttgactgctgcttccgtaaggcttctgaggggggtgactgacagctgtattgccttta  
gcctgcggatagttaatggtgaggagagagagcctctgtgtttcctaagtcagactggct

FIG. 21

72/157

tccctcagttttccacttttagtccaccggactcctagcttgccctccttctgtcttgccagg  
gctgcacagagaacactcacacctgaatccttatcctgcacccttggtaatccgaggcca  
gcaagacttggttctccatactcagttaagccttctcttgggcggagagatctttaacctc  
tcaggaaggcttctctgtccttcttctcctcctctccttctctctcttcagcctccat  
tccaaagccacacttgccacatccccacttgccacatggctcactaattacgaagcacatgt  
gactttaaggacgaggttccctggagctgctgctatcttcagattttgaaaatgtatatgag  
cgttttgccctgcacatgcatctgtgtaccgcggaggtaccaggaaaacctggggggccgga  
agagagaggcggatcccctggagctgggggttacagatagttatgaggcaccatgtgggtg  
ctgagaactgaacccggctcctcaggaagaacagcagtgctctcaatccccgatccacct  
ctccagctcccattccagaatcttgatgtctgggatctagggaggcgggtggaacaaggag  
ctgggatgacccaggaagggtgagccaacagagctctccactcaggcaaggccagagct  
ggagtggctgggggttctatgttttttctctttttttctgtaacctgtgccaaactgga  
agatgcagcggcagagggataagtgggcttgctgctatctctgcaggagcttggtcctg  
caggataggtcctgtgagtggttctgggaccagataggagatggggaggtaggaacattc  
ctgaagaacattccaacaccgtctgagtaatgcttggctccgattctcttctcctggagccc  
cttggaacgcctggcccaggatcactgagcctctgaacctagctccttcaggttctgg  
gccaggtccgtcagagtcccaacaccagcatcaaccctaattgggagtcaggagctgtggg  
acagacatttggtggctgtcttactcagatggcaaggaggcctccāaacacctccctca  
ccataaaggcatctactaccagctctaaaaataaccaaggctgcctccttcttcccagcc  
cctctgcgctgggggtctgggcttcttcccgggacctctgacctcaaaccagctctctaag  
ctctggccgcttcaaattccccaatggggatgggttccctgggtgcctcgggttctcttttgc  
cacctccctgggttgcttctccatttaggggatgccccatgaataccaaatcctggaagtg  
ctacacatggagtcagctttcaggaccctccttcccogatggcttccctgctcctaact  
gcaattttaacactattctctgctttcaccctcaggaaatgcccatggagagctagtcaact  
ccccactggatcctggaggggcagactctggctcaagggtcacctggaggagccggttaagt  
atccgttatcctgtccttattaggggccctaggtctcccgtctctgtccttcccccaa  
atgcagcgtgacagttttctgggggttctcagaaccaaatacactgtaccgttaagccact  
ccctgttgggtgggtgggacaaagagacatctagatggcgtgaggctgcacacatggaag  
tcattttgggcttccctcccaccccccttctattcctggccctaaccctcttctccat  
gggggaacattcaggccatgcccactcaactaagcaaagatgagttcatagtttggggcc  
taaagcctttcaggagagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  
gtgtgtgtgtattctgttctgttttgattgagacaggggtcttgctatgtagctct  
ggggggcttcaagtcagccatcatgccagccacaggagatttttaaaagaaggacac  
caaaatgtcctctttggccgtgtttatggggacatgtgactccgtggacatgctgctgt  
cctgtaacctcagcttccatggctggaagccacctgggacaaggaagctgtgggccattc  
tccctcctgtcccagatgtattctcactctgggttgccttctccagagccctgctcagcct  
cctcctaacttctgctcattgtgttccccagaagagccatgtgtgccttccactgcttcg  
ctgcccatggggaggccggcactgatccgcctggcttacagagaggcaagcgagccccac  
agagcttgtaaagttgctcatgcttgaaatagttcattgggcctagaaacctgaggtccg  
tgggacctaaagggttcagatgattggctgcctcccctcttttctataactactcaggc  
tggagagagttcaattctgtacctcctgacagagggcagccccactgtccaaacctagag  
tggtgcaacagtgaggatgcattcccagaaacctaaaccttccaagtcagggccttccact  
gggggtccccagacttctaggacttctctcaaccaagcaaagatggcttgtggacatgca  
cgggttcccctgatgggtggcctctctccttgccctggctcgggtctgtgagagaatcccagg  
gggactctgctctgcgttaggaagcctgtgaggccaggccaggaagagctttggcagggt  
gtgtgtatttcacaaacagggttgactgcaggatggggatgggtgcacaaagggggaggg  
agaccctggagcagaagtaaggcagcaaggggccgcagaggggaagggaaccttccactggg

FIG. 21

73/157

gctactgcctttctcccagagcagacattttcccataaagcaagaggcactccaaccata  
taagctcatgtttgggcccttgtggagcctgtggctggggaagtgagggccatcttctct  
acacctgctgcagaagggccctgaaagattcttggccagggcccagcccagtgcat  
gggataaaaagggaaagccatcgtgggtggggaacatttaaaaaaaaaatacagcagc  
ctccctggaatctcttgggctagttccagttctggcttctagccaggctaagtggacta  
gcctgagagagaccaagtcagtgagagaggagaggtggctagagggccaaggccagccct  
tctgacatctagctaagagagtcacacactttttagggagccaaagttgggtttcatgttt  
acttcatgagttcagttatgaggccagctcagaggaatttcagaggatcgggcagtttgt  
cagaactgaaggtggaggaaaaagtttgggttctctcaggaatgagggaaaggcagcaag  
atggagagatgtgcaaccaggagctgtggatgcagctgggtgatttgggtgcttgcctggc  
atgtgcaaggccctgggtttgatctatagcaccacagcaaaactagcaagcaagcaaaaca  
agcatacaaaacaaaacagaagcaaggttgcaggtagtacaagagatggccaaagctgtag  
cccaccaatgaccacatcacctcagtgccactgcctcctgcttttccctgccttttc  
atggggctcctgactatggccatagcatatgtccatagcagacataacaacatagccag  
gcaacgtagcaactgctatgtccgtatatgggtgggttatgcatctttaagggtccgttgtc  
acaaacacaagcactgagaacatctctaagttacaactcccacacatatccactccctgc  
aatgagagagccagctgttccctatgagctttctcactaggcagctctacatcctaccggt  
tccctgggcagccaggtggccttggggcctagtggtgctcactgcgttccttctcggtcagat  
cttgaagcctgactcggtgctgggtggctttagaggctgaaggccaggatctcctgcttga  
actggagaagaagcagtgagtaccagcgggggggttggctgaagtccagacaaagaccctc  
tctggagaggatcagtgctttctgggagggggttgggggctgggtgggaagcagcagtg  
gagtgacagggagagtggtggatgtacttggggctacagtggaactgaacccaactgtagg  
aagtacatacctggctcatctcataccctggagcagctggatgggcggctcccgcctggag  
aagagtgagcaccagaacagacattgactgataccttaattcaaaggggaattcttaggca  
aagggaacttccacagatggcagaagaaaaagctgcgaaggtcaaataccaatggtgacag  
ttcatctggtctacctggaatgcctccatctgcgtggaagagaacagagacatctggtct  
acctggaatgcctccatctgcgtggaagggaacagagaggagagggtgggtggggggc  
ttccagtggggccgatcgaagagggcatcttagtgctaaagatcttggatggacctaaa  
gagtcaaagaacgggaggtggttaaaggagaaacacagtggtcgaaggagtgagagggtaa  
gcaagatggaggcgtggaactgcagactccaccaccaaccacaccactcgctcatgtctg  
ctttcttacggacattgcagaaggcactcattcattttattctctccatggaatcagaca  
gactgggaacaaaactgcaaacaccacgccaacacacagttcctgacagtgccacatccct  
tctactaaacgggtcttgtacatgtgtgcacacattgaacgttagcatgtatgagtgcag  
ccagcaaaagcagcagagttgtaattgtgtgtaaggatggctggcgtacaacagcctgggt  
ttttgatggaggacaatgtgaatttggagtcaggattttctgtaaggaagagcaattgaa  
gtcaggcatggaggtagaggcctgtaagtcgggctcctctgggcaggagaacacagagtc  
aagtcaaggactgcctgagttataaagaatgagttcaagaccagcctaggcaacttagca  
aaatccagttctcaaaatataaaggagtgaggaagaggggaaagagaaaaagagggagaagg  
agaaggaggaaaaagaaaagaaagaaaggaagaaagaaagagaaggaaggaagaaaggaag  
gaaggaagataaaagaaagagagagagacagacagacagaaagaaagacagaagacagaaa  
gaaagacagaaagaaagaaagaaaggaaggaaggaaggaaggaaggaaggaaggaaggaag  
tcctgggtggggacaggggaagaagcctgaaaatggaacattactgatagagtcctagctcc  
aggtaaagggacccttccccgggtgggtgaacagaagcagctacaggcaggtggaagtcg  
gaagtacgcatcactgagggacagatggagacatgagatctagcagctctgaggagaactg  
gagggcctgggtccgctcagcaaaaggcaggagtagccctggccctacttggcctctctga  
ggtagtatctcctggcac  
acctgcctgcagcaggctccacagcaccaaaggtccccagcctgtccttctgggtgcttc

FIG. 21

74/157

ccacacttcccatccccttgactaaccacactcaggactcgctccagacctttgtcttct  
ttttgcctgaggttttctctttgaggcacagtgtctcctttctctctgacacactgaatg  
cctccgctagcactcagggaggagagctggaagcaagcggcaggatgcaggtctaaatcc  
agtgactcctaaccctccccctttccagcaagcttctggccccaggatacacagaaacc  
actacaggccagatgggcatccggtagtgtgtcccccaaccacacgggtgagacactttc  
atgggatctgagaacatggggaggggtgcagcctcaatgcctaccctgtcctgttttgc  
tcctaaggaggttgtagtctgccttcaggccccacttcctccttgtcccaccaacttgag  
ctcctactcagatgtgggtatgagcctcttctctctaggtcattgccaatatcacggg  
cgtgtgaggggcttccgggaatcctgggtgggttctcagcacctgctctgggatgaggtaa  
ggcgttgggggagagggcctggggctgcggcagaggtgagcgcttctgttcacacatcc  
ttgctatccactgcaactctagtggccttattgtgtcagcagcaaagtgcagctattatct  
gcaacctcggactcctggggataccaaagacttcccaaccacagagatcttccggatgga  
gcagttgttcacctggagaggggtccagagagacaagaactcccaatacaaaagcaggaat  
ggccagctcttctcatgtccccagagccgggttaaggggtacagactagagtgggactgg  
tccttcaaagcagaggacacaggaatgcgggtcctggacgggggtctcctgttttaattc  
agaggctgaagaatcaagtttgagcagaagtttaactatccccctccccctagccccggggg  
atcacgggtgtgtggccaaggattcgggtcaagcagcctgccccctcaccatttgcctccc  
aggtgaggcgagaggcgcgaggagtccaggtacctggaactgtacatagtggctgacc  
acacctgggtgaggagagacctaagaggtctgctgggtggggggactgctcagcaccgtt  
atgaccagggacctctctgccttctcagttcttgcctcagcatcagaacttgaaccacac  
gagacagcgctcctggaggttgccaattgcgtggaccagggttggtgattagggcgctg  
aaggttgggagtgaaggcagttgaaaagctctggtgagtgcctggttctggttcagcag  
gtctctccctcagattctcaggactctggatatacagttggtgttgaccgggctggaagt  
gtggaccgagcaggatctcagtcgcatcactcaggacgcaaacgaaacgctctgggcttt  
cctacagtggcgccgcgggggtgtggggccaggagaccacacgactccacacaactgctcac  
gtgggtgccatttaca  
caccocgggttgccggatgtccttatttccaactccaccagtcacgccttgcctcatcat  
ccttagggggccgcaccttccagggtaccacgggtgggcctggcacctgtggaggggcatatg  
ccgcgcggagagctccggaggtgtgagcacagtaagctctgctggggacctaggaggaga  
gaggtttgggtagggtttgtgggtcctcagtgaaaccggctcctcaggaccactcggaac  
tccccatcggcacagcagccaccatggcccacgagataggccacagcctgggcctccacc  
atgatcccgagggtgctgctgctgcaggccgatgcagagcaaggaggctgcgtcatggagg  
cagccacagggtacgcaaagaaggtcgttgttacccggacctagacaaagcgactgcgtat  
ctttgggttaccctcagttccctcttttagtaaagataacagcctgctaatagggcctcgg  
tgggaaagatgctgataaggtggagtctcagaatgactgcgggactagggtactccactc  
tgatccttgggtcctccttggggccccggactgcccgatttttagcgtcacagcctcctgtct  
ccagttccttaggtcctctctctctctcgattgctcgcttagcttagaaagaaacagcacat  
cttgagggtgtgtggctaacctcaggggcccagagtgcagaagtagcctgagcttaggag  
gtggctccagcacccgagatggtccccacaccaccaatgacatccatcttccatctgcag  
gcacctttcccgcgctcttcagcgctgcagccgcggccagctgcgcaccttcttccg  
caaagggggcggtccttgcctctccaacacctcggcgcgggggtcctggtgctgccag  
ccgctgcggaaacggcttcttggaaagcaggagaagagtgcgactgcggttctggccaggt  
cagatcgtcatcttctgcctgggggttaagggttaacccatgctcccagcctttccagg  
tctagcttgcctccaccatcacgtgttctgttctctcctttctgacttcagaagtgccag  
accctgctgctttgccacaattgctccctgcgtgcggggggtcaatgtgccacgggtg  
attgctgtgcgaggtgcctggtgaggatactggggagcgccctgtacatcttagctgggc  
tgggacttctagtcccttttctgagtgtgagtttgaccagtgatgggttcacgtagca

FIG. 21

75/157

ctccctcgggctttcagcaaagatctctctgctttcccttaatgggtctctggtgtagtt  
aaagtcgcggggcagcccttgctcgtcctgctgcgactgactgcgatctccccgagttctg  
caccggcacctccccgtattgccccgcagatgtttacctactggatgggtcaccctgcgc  
tgaggggtcgggctattgcctagacggctgggtgtccacgctggagcagcagtgccagca  
gctatgggggctggtgagaccgcacgctgggtcctgggtgccctgaccaataactaaaacct  
gcggttttctactgagggcaagctccaccctgggaactgaggcccgagctgcccgcttct  
tactccccctccccccagggtccaagccggccccagagccatgtttccagcagatgaact  
ccatggggaattcgcaagggaactgtggccaggaccacaagggtagcttcctgccttggtg  
ctcagaggtgaggtgtgatgctgaggggtctgcagctgtaaagttagggcggagcatgcgga  
gggaacactccaagttgttgaccaccttccacttccctccccagggaacgctctgtgtggga  
aactgctgtgccaggaggaggagccgaacccactagtgccgcacatagtgactatggact  
ccacaattctcctagagggccgcgaagtgggttgccgagggggcctttgtgctcccagata  
gtcacctggaccagcttgacttgggtctggtagagccaggcaccggctgtggacctagaa  
tggtgagccctgccacccaacccctcctgggtatttgagtcctccatgccaaagtgttct  
cctcactgccagtgggcacaatgcccataggtgtgccaggacaggcactgtcagaatgc  
tacctcccaggagctggaacgttgcttgactgcctgccataacgggtgggggtgagtagcct  
aaggggtcaggggtgaccttggaggtccttgcctacctgggtgacttttctatcctcatctta  
ggtttgcaatagcaatcgtaactgtcactgtgctgctgggtgggctccacccttctgtga  
caagcctggcttgggtggtagcgtggatagtggccctgcacagctctgcaagtatcccaat  
gggggtgggggcaggcaggaaccacctggggcagtagcctgcttgagactcagcaccctgc  
cctccacagaccgagatgccttcccttggccatgctcctcagcttccctgctgcctctgc  
tccttgggcttgccctagcctgggtgctactaccagctcccaacattcttgtcatcgagg  
gactgtgctgcaggaggggacccctatggaataggtaggttcgggtgctcaggtctctctt  
cctgagcctgcccccatggctcctgcttctcagaactcttcagggtttttagagtgaga  
ggctacagggtgctggggcttttaggaagctagatgggatccttattctctagatgtagtg  
agagctcccagggtgtgggaagaagtccgtgggtgttatcactgccctgaccagcacttg  
gtgggtgtgggtcctttccatgggcttcccttgatttttttggttttatgttttttggttgtg  
ttgggttttttctgctttaaagaaattcaacacagcctgtgctccgcttgtcctgaaga  
cgactctagcttccttgtctcaatatcagccatctcccatggcctttgccctaattactc  
ccttcagggtccctcggtgctagccaaggacccttgggtgctccctataaactggccagct  
atagtgttgctctttctgtgccaggactctgtgcctctgtcccctgtagatcacagtgct  
gtaaacccaattttctggcaggcagaaatgccacctgatagcacacattatctaccgagg  
gcagcctcctcactaggccctcgggcaagtgcaggcacattatgtccctgtctgaacttg  
gaagtgttctagaccaattgaggggaagggtgggttgggagatgtgactagagggc  
acctcaggccagaaacagcaccagcaggccccaggagccagtgagacgggtctggggaaag  
ccaggtagtgcctgggggggtgggggcgggtgtctgcagacagggaacaggtggagtga  
gttgggagggggcacttcagaggggtggcagctgcacaccgttatcgggataggggtgtca  
aggacagtgggaatatctggatgaaccatccaaagaaatggcaagggtctgtgagaaagg  
cccggcagttcagagtctagactggcagagcacggctaaacggagcatgggaaggcacag  
tccccaccaaggagacatctctgcacttcagctctaggtgggcccctcggtgacgcctac  
atctagactgagtgggggtggagtgagtcacctcgggagaaaggatgtccacagccct  
ggaggccctgggacataagtgggtctggacatcttaaggacagaaacaggaatgtggaat  
tcgaagcttgagtggaaatggagaagcaatcccccttggctcatacacgggtcactttcca  
accttctgaactcttatctggtctgtcactggcccctgcagagacataccctgggcagt  
gtgcatccgggtggagtttggctccatcatcactggagagccctcgccccctcgtaagtgt  
gccccttgggacatggagaggggaagcaagggtgggtgcttgccatctgcctccttctaa  
atgctctccttaacacaccttcaatcctctgcctgctcagccccatggacctcttgcca

FIG. 21



76/157

acagcggttcgcaccctccatctcttgacttgctctcagaccctgcgaactctgagcttac  
ctaagaactaccctctgaagcagcctgggtctacagattgagttccagaccctgccttatcc  
ctatgggtatggaagcaccctgaggacctcctgttgcccagtcacctacctctgtctcagt  
ttgttggtccctcctcagatttacaggccttgcatcaataaagaaatgagacatgggcctc  
agagaagctgttgatagagaccatgatgctggaagccctaggggcaggggaaggagac  
actgtgggttcttcttggtccttatagagggaggacaaatgtgccctgccatgtgacttg  
cagtcctcagtttctcagacgcactcttataattcctatgggctgtatgctgagctctta  
ctcagcataggaaccccagagcccgatcatgttgatcccgctgccctgagagctgtgc  
tattctgaaatgtagaatgtatctaataacaataaatccacaagttatatcagtggtgt  
tggtctgtgacctgttaaaaggtctaaagtgtttattaaaaagatatggagatggattac  
tgagaaagaaattagaaacaacatctggacagtgaggagccagcactggggaggaaagg  
gcagacagatctctgagttcaaggccagcctgggtctacagattgagttccaggatagcta  
gggccacacagaggaaaccctgttttgaaaaccaaacagtcaaataataaaaacaaactt  
agtagcaccttgtagacagagagaaaaggttccaggagagctcagaccccagcaccag  
aggtggcaaggcagagctctcaaagcccaggtggcaaaagtgggatctgttagcataaacc  
caaagggcgctgggacagggcagggagtagcctttgatgcaagctcactctggtagggc  
ccctcacccctggatgtctgttagcaagggaatcagtcagtgctcagtcgtttaacatc  
tgtgagaggggaaaggctgctgcagacatggcctgagcagcatctggattcgaacatttg  
cactttagggcctgctctctcccctgggtggggcactcgccattcattgcctttacgaca  
gctgtgagagagaggttgctgcaagtgtgtatggctgggttagctcgagcccaccaggca  
atcatgatttccctcgacatctagtcattaacaaacacaggttcttttactacaatttta  
actaccaatattaactaataatgtgtataaaatatataaacacagtagacatatataagta  
catagatgttagtagacatatataatattatatatatataattgtaattaatattactt  
aattttattttatcatttgatattattttactgttagttataacaatgtgcataatatat  
gtgtaatatataaatataattttatttttaattattattatatataaatttaattaatatta  
ttatacctatatatttagtagacatatacataggttacagaatggctacaaaagtgccagga  
gccatcaaggagaagctaaaagccagcaagtgatcttctgagacgggttctgccatggac  
tgtacaattagtgatggatttgcttctgtaggcaaggacgaggagatttcattttaggaa  
agattcctgctattaatatgcttttctggtattattaaatatataacaatcactagg  
tattagcccaccgttttgaacagaatgttctgcagaacaatgaagatgtactctcttgta  
atgatgctatatagacaaatagattatttcttttttaaaaaagaaaaagagccgggga  
tggtggcacatgcctttaatcccagcacttgaggaggcagaggcagcagatttctgagtt  
caaggccagcctggactacagagtgagttccaggacagccagggtactcagagaaactc  
tgtcttggaaaaaaaaaaagaggaagaaagaaaaagattttattttattttatatcat  
atgagtagacccatcagacacacaagaagagggccaccagaccccattacagatggttgga  
gccaccatgtgggtgctgggaattgaactcaggacctctggaagaacagttgggtgctctt  
aaccctgagccatctctccagcccaaataagatgatttcttaattcttaaggatgatcct  
ataagaattcctaaacttacattagtaattattaagctcttttacaataggacttctatt  
aagtccttctctaatatgaaaacttcaataagaactctgccagtcctcaagtgatgagtg  
tagttgcttctgagatagcaagtaggcatcaacaacttagagcacattctaggaggtgt  
aaaaccattaaccagtggtcttaaaaaggggaactaacaataggctataggtgcaaggaca  
gaagataaaatattgactaggtttatcaatacaaaatttaccacaaaagttatgttttt  
gacttttcataaaaactctttatgaacctgtagaactgggtgaaagatgacgaatgcttag  
ccagataaattactcctaataagatatgcatgtgaatattctgtgctgtaaaacttatttatg  
tttgaacttccagtgaaactttgttttaaaaaaggggggggttgaaaaagccatgtgatc  
tattctcctagaaagggtagagaagactaagaagattacattggagatgtaaccttgga  
gagaaagctttgggagcaagagcatagagagcaaggccattgtggcatcagagcaggagg

FIG. 21

77/157

agagagcaagattagaaggagatgcagagtggaataacttagaaactataaggcaacata  
aaaaattaagagagccatatgcagaatgcagagggaaagagaaaaaaaaaaaaaaga  
agctgcaggagagcagaaggagcaggcaggcttctcctgaccatggggtagaacagggc  
ttttcttaataccaaggcaggcttagtcttaaggataataaagcttttctttcttacaga  
cttggttttaattcatttagcaataaaagtgtaaaagtgttttctttccctatgcaataa  
agattggagcttatttttcagccagaatgagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  
tttgcttcataacacacataagt  
gtgtgtgtgtgtgttaagtgtgcaattatcagatggcatggaagctgggctcaattggttc  
aaatggggacttgtgagggtatatgcatgaatctgtatatgaattcatgtgagcttat  
atatttgcttgtgtaaaagttttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  
agaggtttattgcttcaaacttccccctagcctgacagtcgagaggcatctggacaagag  
agaaaaggctctagccattaatccttttcttagatccattttcttagagaactttcttag  
gaaactgttttagagagaacatagaaacagggtgaaatcacttgtcaaactgtcccctttt  
cttcttaaggacttctactagcagactgggagtttagagctgcacagtccctgaggagata  
gaacaaaggctgctttactgaatccccctgctgttttaagatgaggttctaaaggagattg  
cagtttctgacccccaaaaggaaactcaggcaggctcagctacagtatcaaagtgtctaaa  
cttaagatagggatatgttttattattaaacagctaccctaaatatctcataagatcaag  
cttaccctcggtgacacttccccctctgttgctcaagaggaaccaaagcagaaagaaccgc  
cagggtggtcctgggcacaaatgggttaaagatgttgtagcatggggaaatgaagagat  
ggctcagctattaagagaatatcttactcttcagaggaccagtggtcaattcccagcaa  
acatatcaggtgccacaccatcacttgtagctccagctgcagatctgctacatctggcct  
ccataggcacccacacacagggtggcacccacaattaaaaaaataagataaatctaaaaca  
gcaaagttaaagcatgagctgaaactagtaaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  
ggtttggtccctacctgttagaaatagtctcagtatcacacaaaggaacacccaagcgaa  
gcaaaagctcccagcaagacaaaactacagtcttcattgagagtgtgcacgctgaagacc  
gagcacactgggtgcaaaatgtacttggattctgtttgcttgttttgttccagacagggt  
ttctctgcataacagccttggctgtccctgaaattcactgtgttagaccagggtggcctca  
aaccctcagagatccgttcacccacgcttatctaggcttcagtctcacctgtgagatgg  
cctgaaagtgttagaaccgcgaggatctatttctgacagactggctggcatcttttcc  
ttctctcagcatgagattcctggggcggtcccatctcagcatcaagcatggtagcagagt  
tggaacctgagggtgagggtcagactcagaccataaactggaagcagagagaactgga  
gattgtgggaggctttgaaacctcagctcctgccccagcaaataccttccagcaaggcca  
cacctcttaaacctccccaaacagggtcaccaactggggacctaataattcaaatgccac  
aaatatgggagacatgacatccaaaccgcccaggacagggtgtatacctccatgcttgggtt  
ccgtagtaagaaacactaaacattagcctttcctaataaacactgatataaagccctgct  
attctcgatgtttttctctgttctgtctcctccttctccacctgcttctctgttctctga  
cctcttctgtgtcacagatagccctgccatgtccatctgccagccatgttctgtctactt  
gcctctctctctgtctctggactcttctagatgcctctggctgttctttctcatatctaca

FIG. 21

78/157

>mouse Gene 216\_cDNA  
AAAGGCACTCCCAGCCATGGGCTCGAGGTGCGGGAGACCCGGGGGGTCTCCG  
GTGCTGCTATTGCTGCCGCTGTTGCTGCCCTCGTGTCCGCTGCGGAGCGCTCG  
GATGTTTCCAGGAAATGCCCATGGAGAGCTAGTCACTCCCCACTGGATCCTGG  
AGGGCAGACTCTGGCTCAAGGTCACCCTGGAGGAGCCGATCTTGAAGCCTGAC  
TCGGTGCTGGTGGCTTTAGAGGCTGAAGGCCAGGATCTCTGCTTGAAGTGA  
GAAGAAGCACAGCTTCTGGCCCCAGGATACACAGAAACCCACTACAGGCCAG  
ATGGGCATCCGGTAGTGCTGTCCCCCAACCACACGGATCATTGCCAATATCAC  
GGGCGTGAGGGGGCTTCCGGGAATCCTGGGTGGTTCTCAGCACCTGCTCTG  
GGATGAGTGGCCTTATTGTGCTCAGCAGCAAAGTCAGTATTATCTGCAACCTC  
GGACTCCTGGGGATACCAAAGACTTCCCAACCCACGAGATCTTCCGGATGGAG  
CAGTTGTTACCTGGAGAGGGGTCCAGAGAGACAAGAACTCCCAATACAAAGC  
AGGAATGGCCAGTCTTCCCTCATGTCCCCCAGAGCCGGGTGAGGCGAGAGGCG  
CGCAGGAGTCCCAGGTACCTGGAAGTGTACATAGTGGCTGACCACACCCTGAA  
CTTGAACCACACGAGACAGCGCTCCTGGAGGTTGCCAATGCGTGGACCAGA  
TTCTCAGGACTCTGGATATACAGTTGGTGTGACCGGGCTGGAAGTGTGGACC  
GAGCAGGATCTCAGTCGCATCACTCAGGACGCAAACGAAACGCTCTGGGCTTT  
CCTACAGTGGCGCCGCGGGGTGTGGGCCAGGAGACCACAGACTCCACACAA  
CTGCTCACGGGCCCGACCTTCCAGGGTACCACGGTGGGCTGGCACCTGTGG  
AGGGCATATGCCGCGCGGAGAGCTCCGGAGGTGTGAGCACAGACCACTCGGA  
ACTCCCCATCGGCACAGCAGCCACCATGGCCCCACGATAGGCCACAGCCTGG  
GCCTCCACCATGATCCCGAGGGCTGCTGCGTGCAGGCCGATGCAGAGCAAGG  
AGGCTGCGTCATGGAGGCAGCCACAGGGCACCCCTTCCCGCGCTCTTCAGCG  
CCTGCAGCCGCCGCCAGCTGCGCACCTTCTTCCGCAAAGGGGGCGGTCTTGC  
CTCTCCAACACCTCGGCGCCGGGGCTCCTGGTGCTGCCAGCCGCTGCGGAAA  
CGGCTTCTTGAAGCAGGAGAAGAGTGCAGTGCAGTCTGCGCCAGAAAGTGC  
CCGGACCCCTGCTGCTTTGCCACAATTGCTCCCTGCGTGCGGGGGCTCAATG  
TGCCCCACGGTGATTGCTGTGCGAAGTGCCTGTTAAAGTCCGCGGGCACGCCTT  
GTCGTCTGCTGCGACTGACTGCGATCTCCCCAGTTCTGCACCGGCACCTCC  
CCGTATTGCCCCGAGATGTTTACCTACTGGATGGCTCACCTGCGCTGAGGG  
TCGCGGCTATTGCTAGACGGCTGGTGTCCACGCTGGAGCAGCAGTGCCAGC  
AGCTATGGGGGCTGGGTCCAAGCCGGCCCCAGAGCCATGTTTCCAGCAGAT  
GAATCCATGGGGAAATTCGAAGGGAAGTGTGGCCAGGACCACAAGGGTAGC  
TTCCTGCCTTGTGCTCAGAGGGACGCTCTGTGTGGGAAACTGCTGTGCCAGGG  
AGGGGAGCCGAACCCACTAGTGCCGCACATAGTACTATGGACTCCACAATTC  
TCCTAGAGGGGCCGGAAGTGGTTTGCCGAGGGGCCTTTGTGCTCCAGATAGT  
CACCTGGACCAGCTTGACTTGGGTCTGGTAGAGCCAGGCACCGGCTGTGGACC  
TAGAATGGTGTGCCAGGACAGGCACTGTGAGAATGCTACCTCCAGGAGCTGG  
AACGTTGCTTGACTGCCTGCCATAACGGTGGGGTTTGCAATAGCAATCGTAACT  
GTCAGTGTGCTGCTGGCTGGGCTCCACCCTTCTGTGACAAGCCTGGCTTGGGT  
GGTAGCGTGGATAGTGGCCCTGCACAGTCTGCAAACCGAGATGCCTTCCCCTT  
GGCCATGCTCCTCAGCTTCCCTGCTGCCTCTGCTCCCTGGGGCTGGCCTAGCCT  
GGTGCTACTACCAGCTCCCAACATTCTGTCATCGAAGGGGACTGTGCTGCAGG  
AGGGACCCCTATGGAATAGAGACATAACCCCTGGGCAGTGTGCATCCGGTGGA  
GTTTGGCTCCATCACTCACTGGAGAGCCCTCGCCCCCTCCCCATGGACCTCTTG  
CCAACAGCGTTCGCACCCTCCATCTCTTGACTTGCTCTCAGACCCTGCCAACTC  
TGAGCTTACCTAAGAACTACCCTCTGAAGCAGCCTGGTCTACAGATTGAGTTCC  
AGACCTGCCCTATCCCTATGGTATGGAAGCACCCCTGAGGACCTCCTGTTGCCCA  
GTCACCTACCTCTGTCTCAGTTTGTGTCCCCTCCTCAGATTTACAGGCTTGCA  
CAATAAAGAAATGAGACATGGGCCTCAGAGAASCTGTTGTATAGAGACCATG  
ATGCTGGAACCCCTAGGGGCAGGGAAGGGAGACACTGTGGTTCTTCTTGGGTC  
CTTATAGAGGGAGGACAAATGTGCCCTGCCATGTGACTGCAGTCTCAGTTTC  
TCAGACGCACTCTTATAATTCCCTATGGGCTGTATGCTGAGCTCTTACTCAGCATA  
GGAACCCACAGAGCCGATCATGTTGTATCCCSCTGCCCTGAGAGCTGTGCTAT  
TCTGAAATGTTAGAATGTATCTAATAACAATAAATCCACAAGTTATATCAGHAAA  
AAAAAAA

FIG. 22

79/157

&gt;mouse Gene 216\_protein

MGSRCGRPGGSPVLLLLPLLLLPSCPLRSARMFPGNAHGELVTPHWILEGRLWLKVTLEEPILKPDSVLVALEAEGQDLLL  
ELEKHKHLLAPGYTETHYRDPDGHFVVLSPNHTDHCQYHGRVGRFRESWVVLSTCSGMSGLIVLSSKVSYYLQPRTPGDTK  
DFPTHEIFRMEQLFTWRGVQRDKNSQYKAGMASLPHVQSRVRREARRSPRYLELYIVADHTLNLNHTRQRLLEVANCVD  
QILRTLDIQLVLTGLEVWTEQDLSRITQDANETLWAFLOWRRGVWARRPHDSTQLLTGRTFQGTTVGLAPVEGICRAESS  
GGVSTDHSELPIGTAATMAHEIGHSLGLHHDPEGCCVQADAEQGGCVMEAATGHPFPRVFSACSRRLRTFFRKGGGPCL  
SNTSAPGLLVLPSSRCNGFLEAGEECDGSGQKCPDPCCFAHNCSLRAGAQCAHGDCCAKCLLKSAGTFCRPAATDCDLP  
EFCTGTSPYCPADVLLDGSPCAEGRGYCLDGWCPTLEQQCQQLWGP GSKPAPEPCFQOMNSMGN SQGNCGQDHKGSFLP  
CAQRDALCGKLLCQGGEPNPLVPHIVTMDSTILLEGREVVCRGAFVLPDSHLDQLDLGLVEPGTGCGPRMVCQDRHCQNA  
TSQELERCLTACHNGGVCNSNRNCHCAAGWAPPFCDKPGLGGSVD SGPAQSANRDAFPLAMLLSFLPLLPAGLAWCYY  
QLPTTFCHRRGLCCRRDPLWNRDIPLGSVHPVEFGSIITGEPSPPPPWTSCQQRSHPPSLDLLSDPANSELT

FIG. 22

80/157

mGene216 1 MGSRCGRPGGSPVLLLLPLLLPSCPLRSARMFPGNAHGELVTPHWILEGR 50  
 || | | | | . | . | | | | | | | | . | . | . | : | | | | : | : | .  
 hGene216 1 MGWRPRRRARGTPLLILL.LLLLLWVPFGAGVLQGHIPGQPVTTPHWVLDGQ 49

mGene216 51 LWLKVTLEEPILKPDSVLVALEAEGQDLLLELEKKHKLLAPGYTETHYRP 100  
 | | . | | | : | | | | | | | | | | : | | | | | | | | | |  
 hGene216 50 PWRTVSLEEPVSKPDMGLVALEAEGQELLELEKNHRLAPGYIETHYGP 99

mGene216 101 DGHPVVLSPNHT.....DHCQYHGRVGRGFRESWVVLSTCSGM 137  
 || | | | . | | | | | | | | | | | | | | : | | | | | | | | | |  
 hGene216 100 DGQPVV LAPNHTVRCFHGLWDAPPEDHCHYQGRVGRGFPSWVVLCTCSGM 149

mGene216 138 SGLIVLSSKVSYYLQPRTPGDTKDFPTHEIFRMEQLFTWRGVQRDKNSQY 187  
 || | | | | | | | | . | | | | | | | | | | | | | | : | :  
 hGene216 150 SGLITLSRNASYLWPWPPRGSKDFSTHEIFRMEQLLTWKGTGCHRDPGN 199

mGene216 188 KAGMASLPHVPQSRVRREARRSPRYLELYIVADHTL.....NLNHTRQR 231  
 || | | | | | | | | | | . : | | | | | | | | | | | | : | |  
 hGene216 200 KAGMTSLPGGPQSRGRREARRTRKYLELYIVADHTLFLTRHRNLNHTKQR 249

mGene216 232 LLEVANCVDQILRTLDIQLVLTGLEVWTEQDLSRITQDANETLWAFLOWR 281  
 || | | | | | | : | | | | | . | | | | | | | | | | : | | | | | | | | | |  
 hGene216 250 LLEVANYVDQLLRTLDIQVALTGLEVWTERDRSRVTQDANATLWAFLOWR 299

mGene216 282 RGVWARRPHDSTQLLTGRTFQGTTVGLAPVEGICRAESSGGVSTDHSELP 331  
 || . | | . | | | | | | | | | | | | | | . | | | | | | | | | |  
 hGene216 300 RGLWAQRPHDSAQLLTGRAFGQATVGLAPVEGMCAESSGGVSTDHSELP 349

mGene216 332 IGTAATMAHEIGHSLGLHHDPEGCCVQADAEQGGCVMEAATGHFPFRVFS 381  
 || | | | | | | | | | | : | | | | : | | | | | | | | | | | | | | | | |  
 hGene216 350 IGAAATMAHEIGHSLGLSHDPGCCVEAAESGGCVMAAATGHFPFRVFS 399

mGene216 382 ACSRRQLRTFFRKGGGPCLSNTPAGLLVLPSCGNGFLEAGEECD CGSG 431  
 || | | | | | | | | | | | | | | | | | | | | . | | | | . | | | | | | | | | |  
 hGene216 400 ACSRRQLRAFFRKGGGACLSNAPDPGLPVPPALCGNGFVEAGEECD CGPG 449

mGene216 432 QKCPDPCCFAHNCSLRAGAQCAHGDCCAKCLLKSAGTPCRPAATDCDLPE 481  
 | . | | | | | | | | | | | | | | | | : | | | | | | | | | | | | | | | |  
 hGene216 450 QECRDLCFAHNCSLRPGAQCAHGDCCVRCCLKPAGALCRQAMGDCDLPE 499

mGene216 482 FCTGTSPYCPADVYLLDGSPCAEGRGYCLDGWCPTLEQQCQQLWGPGSKP 531  
 || | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 hGene216 500 FCTGTSSHCPDPVYLLDGSPCARGSGYCWDGACPTLEQQCQQLWGPGSHP 549

mGene216 532 APEPCFQOMNSMGNSQGNCGQDHKGSFLPCAQRDALCGKLLCQGGEPNPL 581  
 || | | | | . | | | . | | | | | | . | | | | | | | | | | | | | | | . | |  
 hGene216 550 APEACFQVVNSAGDAHNGCGQDSEGHFLPCAQRDALCGKLLCQGGKPSLL 599

FIG. 23

81/157

```
mGene216 582 VPHIVTMDSTILLEGREVVCRGAFVLPD.SHLDQLDLGLVEPGTGCGPRMV 631  
|||.|.|||:|:|.||| |||. | | | | | | | | | | | | | | | | | |  
hGene216 600 APHMVPVDSTVHLDGQEVTCR.GALALPSAQDLLGLGLVEPGTQC.GPRMV 649  
|||.|.|||:|:|.||| |||. | | | | | | | | | | | | | | | | | |  
mGene216 632 CQDRHCQNATSQELERCLTACHNGGVCSN.NRNCHCAAGWAPPFCDKPGLG 681  
|||.|.|||:|:|.||| |||. | | | | | | | | | | | | | | | | | |  
hGene216 650 CQSRRCKRNAFQELQRCLTACHSHGVCNS.NHNCHCAPGWAPPFCDKPGFG 699  
|||.|.|||:|:|.||| |||. | | | | | | | | | | | | | | | | | |  
mGene216 682 GSVDSGPAQSANRDAFFLAMLLSFLPLLP.GAGLAWCYYQLP.TFCHRRG 730  
|||.|||||.|. | | | | | | | | | | | | | | | | | | | | | |  
hGene216 700 GSMSDSGPVQAENHDTFLLAMLLSVLLPL.LPGAGLAWCCYRLPGAHLQRCS 749  
|||.|||||.|. | | | | | | | | | | | | | | | | | | | | | |  
mGene216 731 LCCRDPDW.....NRDIPLGSVHPVEFGSIITGEPSPPPPWTSCQOR 773  
||||| | | | | | | | | | | | | | | | | : | | | | :  
hGene216 750 WGCRRDPACSGPKDGPHRDHPLGGVHPMELGPTATGQPWPDPENSHEPS 799  
|||. | . ||| :.  
mGene216 774 SHP..PSLDLLSDPANSELT..... 791  
||| | . ||| :.  
hGene216 800 SHPEKPLPAVSPDPQADQVQMPSCLW 826
```

FIG. 23

82/157

```

      10                      30                      50
CGGGCACGGGTCGGCCGCAATCCAGCCTGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG
-----+-----+-----+-----+-----+-----+-----+
      70                      90                      110
AGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCCGAGAGCTCGGGGGACCCCGTTGC
-----+-----+-----+-----+-----+-----+-----+
                        MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

      130                      150                      170
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGTGCTTCAAGGAC
-----+-----+-----+-----+-----+-----+-----+
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

      190                      210                      230
ATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCTGGCGCACCG
-----+-----+-----+-----+-----+-----+-----+
isIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThrV

      250                      270                      290
TCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAG
-----+-----+-----+-----+-----+-----+-----+
alSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluG

      310                      330                      350
GCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATAG
-----+-----+-----+-----+-----+-----+-----+
lyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIleG

      370                      390                      410
AAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGCTGGCCCCCAACCACACGGTGAGAT
-----+-----+-----+-----+-----+-----+-----+
luThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrValArgC

      430                      450                      470
GCTTCCATGGGCTCTGGGATGCACCGCCAGAGGATCATTGCCACTACCAAGGGCGAGTAA
-----+-----+-----+-----+-----+-----+-----+
ysPheHisGlyLeuTrpAspAlaProProGluAspHisCysHisTyrGlnGlyArgValA

```

FIG. 24

83/157

```

      490              510              530
GGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCCTGATCA
-----+-----+-----+-----+-----+-----+-----+
rgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyLeuIleT

      550              570              590
CCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACT
-----+-----+-----+-----+-----+-----+-----+
hrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerLysAspP

      610              630              650
TCTCAACCCACGAGATCTTTTCGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGGCC
-----+-----+-----+-----+-----+-----+-----+
heSerThrHisGluIlePheArgMetGluGlnLeuLeuThrTrpLysGlyThrCysGlyH

      670              690              710
ACAGGGATCCTGGGAACAAAGCGGGCATGACCAGCCTTCCTGGTGGTCCCCAGAGCAGGG
-----+-----+-----+-----+-----+-----+-----+
isArgAspProGlyAsnLysAlaGlyMetThrSerLeuProGlyGlyProGlnSerArgG

      730              750              770
GCAGGCGGAGAAGCGCGCAGGACCCGGAAGTACCTGGAACCTGTACATTGTGGCAGACCACA
-----+-----+-----+-----+-----+-----+-----+
lyArgArgGluAlaArgArgThrArgLysTyrLeuGluLeuTyrIleValAlaAspHisT

      790              810              830
CCCTGTTCTTGACTCGGCACCGAAACTTGAACCACACCAAACAGCGTCTCCTGGAAGTCG
-----+-----+-----+-----+-----+-----+-----+
hrLeuPheLeuThrArgHisArgAsnLeuAsnHisThrLysGlnArgLeuLeuGluValA

      850              870              890
CCAACTACGTGGACCAGCTTCTCAGGACTCTGGACATTTCAGGTGGCGCTGACCGGCCTGG
-----+-----+-----+-----+-----+-----+-----+
laAsnTyrValAspGlnLeuLeuArgThrLeuAspIleGlnValAlaLeuThrGlyLeuG

      910              930              950
AGGTGTGGACCGAGCGGGACCGCAGCCCGCTCACGCAGGACGCCAACGCCACGCTCTGGG
-----+-----+-----+-----+-----+-----+-----+
luValTrpThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrLeuTrpA

      970              990             1010

```

FIG. 24



84/157

```

CCTTCCTGCAGTGGCGCCGGGGCTGTGGGCGCAGCGGCCCCACGACTCCGCGCAGCTGC
-----+-----+-----+-----+-----+-----+
laPheLeuGlnTrpArgArgGlyLeuTrpAlaGlnArgProHisAspSerAlaGlnLeuL

1030      1050      1070
TCACGGGCGCGCCTTCCAGGGCGCCACAGTGGGCGCTGGCGCCCGTCGAGGGGCATGTGCC
-----+-----+-----+-----+-----+-----+
euThrGlyArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyMetCysA

1090      1110      1130
GCGCCGAGAGCTCGGGAGGCGTGAGCACGGACCACTCGGAGCTCCCCATCGGCGCCCGCAG
-----+-----+-----+-----+-----+-----+
rgAlaGluSerSerGlyGlyValSerThrAspHisSerGluLeuProIleGlyAlaAlaA

1150      1170      1190
CCACCATGGCCCATGAGATCGGCCACAGCCTCGGCCTCAGCCACGACCCCGACGGCTGCT
-----+-----+-----+-----+-----+-----+
laThrMetAlaHisGluIleGlyHisSerLeuGlyLeuSerHisAspProAspGlyCysC

1210      1230      1250
GCGTGAGGCTGCGGCCGAGTCCGAGGCTGCGTCATGGCTGCGGCCACCGGGCACCCGT
-----+-----+-----+-----+-----+-----+
ysValGluAlaAlaAlaGluSerGlyGlyCysValMetAlaAlaAlaThrGlyHisProp

1270      1290      1310
TTCCGCGCGTGTTCAGCGCCTGCAGCCGCCGCCAGCTGCGCGCCTTCTTCCGCAAGGGGG
-----+-----+-----+-----+-----+-----+
heProArgValPheSerAlaCysSerArgArgGlnLeuArgAlaPhePheArgLysGlyG

1330      1350      1370
GCGGCGCTTGCTCTCCAATGCCCCGGACCCCGGACTCCCGGTGCCGCCGGCGCTCTGCG
-----+-----+-----+-----+-----+-----+
lyGlyAlaCysLeuSerAsnAlaProAspProGlyLeuProValProProAlaLeuCysG

1390      1410      1430
GGAACGGCTTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGGCCAGGAGTGCCGCG
-----+-----+-----+-----+-----+-----+
lyAsnGlyPheValGluAlaGlyGluGluCysAspCysGlyProGlyGlnGluCysArgA

1450      1470      1490

```

FIG. 24

85/157

```

ACCTCTGCTGCTTTGCTCACAACCTGCTCGCTGCGCCCGGGGGCCCAGTGCGCCACACGGGG
-----+-----+-----+-----+-----+-----+
spLeuCysCysPheAlaHisAsnCysSerLeuArgProGlyAlaGlnCysAlaHisGlyA

      1510              1530              1550
ACTGCTGCGTGCGCTGCCTGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCATGGGTG
-----+-----+-----+-----+-----+-----+
spCysCysValArgCysLeuLeuLysProAlaGlyAlaLeuCysArgGlnAlaMetGlyA

      1570              1590              1610
ACTGTGACCTCCCTGAGTTTTCACGCGGCACCTCCTCCCAGTGTCCCCAGACGTTTACC
-----+-----+-----+-----+-----+-----+
spCysAspLeuProGluPheCysThrGlyThrSerSerHisCysProProAspValTyrL

      1630              1650              1670
TACTGGACGGCTCACCTGTGCCAGGGGCAGTGGCTACTGCTGGGATGGCGCATGTCCCA
-----+-----+-----+-----+-----+-----+
euLeuAspGlySerProCysAlaArgGlySerGlyTyrCysTrpAspGlyAlaCysProT

      1690              1710              1730
CGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCCTGGCTCCCACCCAGCTCCCAGGGCCT
-----+-----+-----+-----+-----+-----+
hrLeuGluGlnGlnCysGlnGlnLeuTrpGlyProGlySerHisProAlaProGluAlaC

      1750              1770              1790
GTTTCCAGGTGGTGAACCTCTGCGGGAGATGCTCATGGAAACTGCGGCCAGGACAGCGAGG
-----+-----+-----+-----+-----+-----+
ysPheGlnValValAsnSerAlaGlyAspAlaHisGlyAsnCysGlyGlnAspSerGluG

      1810              1830              1850
GCCACTTCCTGCCCTGTGCAGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTG
-----+-----+-----+-----+-----+-----+
lyHisPheLeuProCysAlaGlyArgAspAlaLeuCysGlyLysLeuGlnCysGlnGlyG

      1870              1890              1910
GAAAGCCCAGCCTGCTCGCACCGCACATGGTGCCAGTGGACTCTACCGTTCACCTAGATG
-----+-----+-----+-----+-----+-----+
lyLysProSerLeuLeuAlaProHisMetValProValAspSerThrValHisLeuAspG

      1930              1950              1970
GCCAGGAAGTGACTTGTCTGGGGAGCCTTGGCACTCCCCAGTGCCCAGCTGGACCTGCTTG

```

FIG. 24

86/157

```

-----+-----+-----+-----+-----+
lyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLeuAspLeuLeuG

      1990              2010              2030
GCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAAATGGTGTGCCAGAGCAGGC
-----+-----+-----+-----+
lyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCysGlnSerArgA

      2050              2070              2090
GCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTGCCACAGCCACG
-----+-----+-----+-----+
rgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCysHisSerHisG

      2110              2130              2150
GGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCCTTCTGTG
-----+-----+-----+-----+
lyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaProProPheCysA

      2170              2190              2210
ACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGAAAACCATGACA
-----+-----+-----+-----+
spLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGluAsnHisAspT

      2230              2250              2270
CCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGGGGGCCGGCCTGG
-----+-----+-----+-----+
hrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGlyAlaGlyLeuA

      2290              2310              2330
CCTGGTGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCTGCAGAA
-----+-----+-----+-----+
laTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyCysArgA

      2350              2370              2390
GGGACCCTGCGTGAGTGGCCCCAAAGATGGCCACACAGGGACCACCCCTGGGCGGCG
-----+-----+-----+-----+
rgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuGlyGlyV

      2410              2430              2450
TTCACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCTGGACCCTGAGA
-----+-----+-----+-----+

```

FIG. 24

87/157

alHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspProGluA

2470                      2490                      2510  
 ACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCCCC  
 -----+-----+-----+-----+-----+-----+  
 snSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProAspProG

2530                      2550                      2570  
 AAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAATGA  
 -----+-----+-----+-----+-----+-----+  
 lnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

2590                      2610                      2630  
 ACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAAGTTGAACTGCAGGGGCAG  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2650                      2670                      2690  
 AGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTTCTTCCCCGAGT  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2710                      2730                      2750  
 GGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACATAGAAAGTTCTTGAGGGCTGG  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2770                      2790                      2810  
 AGAACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGGTCA  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2830                      2850                      2870  
 CACAGCCCCTGACCTCCCTCACCAGTGGAGGCTGGGTAGTGCTGGCCATCCCCAAAAGGGC  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2890                      2910                      2930  
 TCTGTCTCTGGGAGTCTGGTGTGTCTCTACATGCAATTTCCACGGACCCAGCTCTGTGGA  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

2950                      2970                      2990  
 GGGCATGACTGCTGGCCAGAAGCTAGTGGTCTGGGGCCCTATGGTTCGACTGAGTCCAC  
 -----+-----+-----+-----+-----+-----+  
 -----+-----+-----+-----+-----+-----+

FIG. 24

88/157

3010 3030 3050  
ACTCCCCCTGCAGCCTGGCTGGCCTCTGCAAAACAAACATAATTTTGGGGACCTTCCTTCCT  
-----+-----+-----+-----+-----+-----+-----+  
3070 3090 3110  
GTTTCTTCCCACCCTGTCTTCTCCCCTAGGTGGTTTCCTGAGCCCCCACCCECAATCCCAG  
-----+-----+-----+-----+-----+-----+-----+  
3130 3150 3170  
TGCTACACCTGAGGTTCTGGAGCTCAGAATCTGACAGCCTCTCCCCCATTCCTGTGTGTGT  
-----+-----+-----+-----+-----+-----+-----+  
3190 3210 3230  
CGGGGGGACAGAGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAAAAGAAAGACA  
-----+-----+-----+-----+-----+-----+-----+  
3250 3270 3290  
TGTTGGCTATAGGCGTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAAGCCGGGGTA  
-----+-----+-----+-----+-----+-----+-----+  
3310 3330 3350  
GGAGGATCACCAGAGGCCAGCAGGTCCACACCAGCCTGGGCAACACAGCAAGACACCCGA  
-----+-----+-----+-----+-----+-----+-----+  
3370 3390 3410  
TCTACAGAAAAATTTTAAAATTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGCTGC  
-----+-----+-----+-----+-----+-----+-----+  
3430 3450 3470  
TCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTGAGTTCAACACTGCAGTGAGCTATGG  
-----+-----+-----+-----+-----+-----+-----+  
3490 3510 3530  
TGGCACCCTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCTAAAATAAATTTT  
-----+-----+-----+-----+-----+-----+-----+  
3550 3570 3590

FIG. 24

89/157

AAAAAGACATAA  
-----+-----+-----+-----+-----+-----+

3610

AAAAAAAAAAAAAAAAAAAAAAAAAAAA  
-----+-----+

FIG. 24

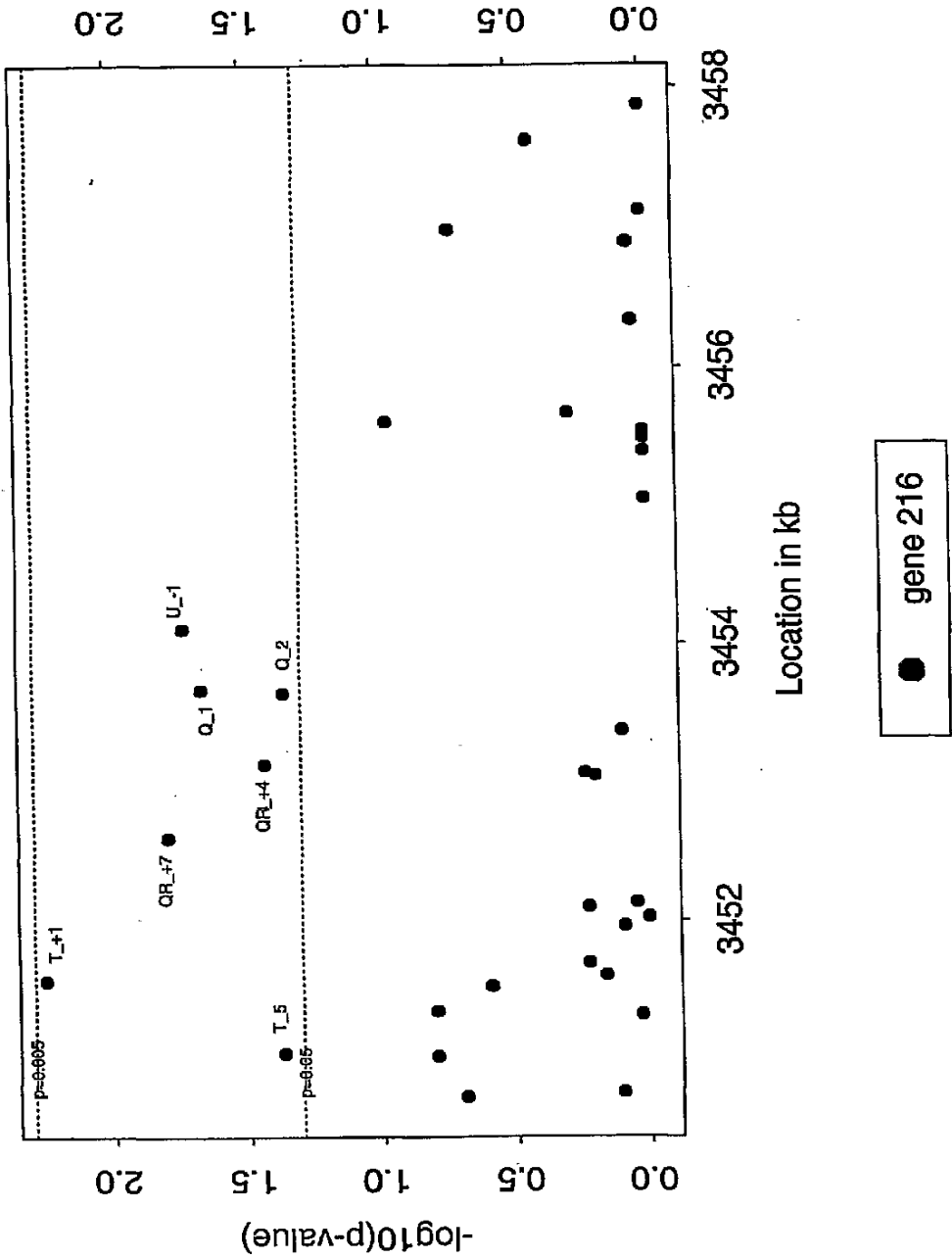


FIG. 25

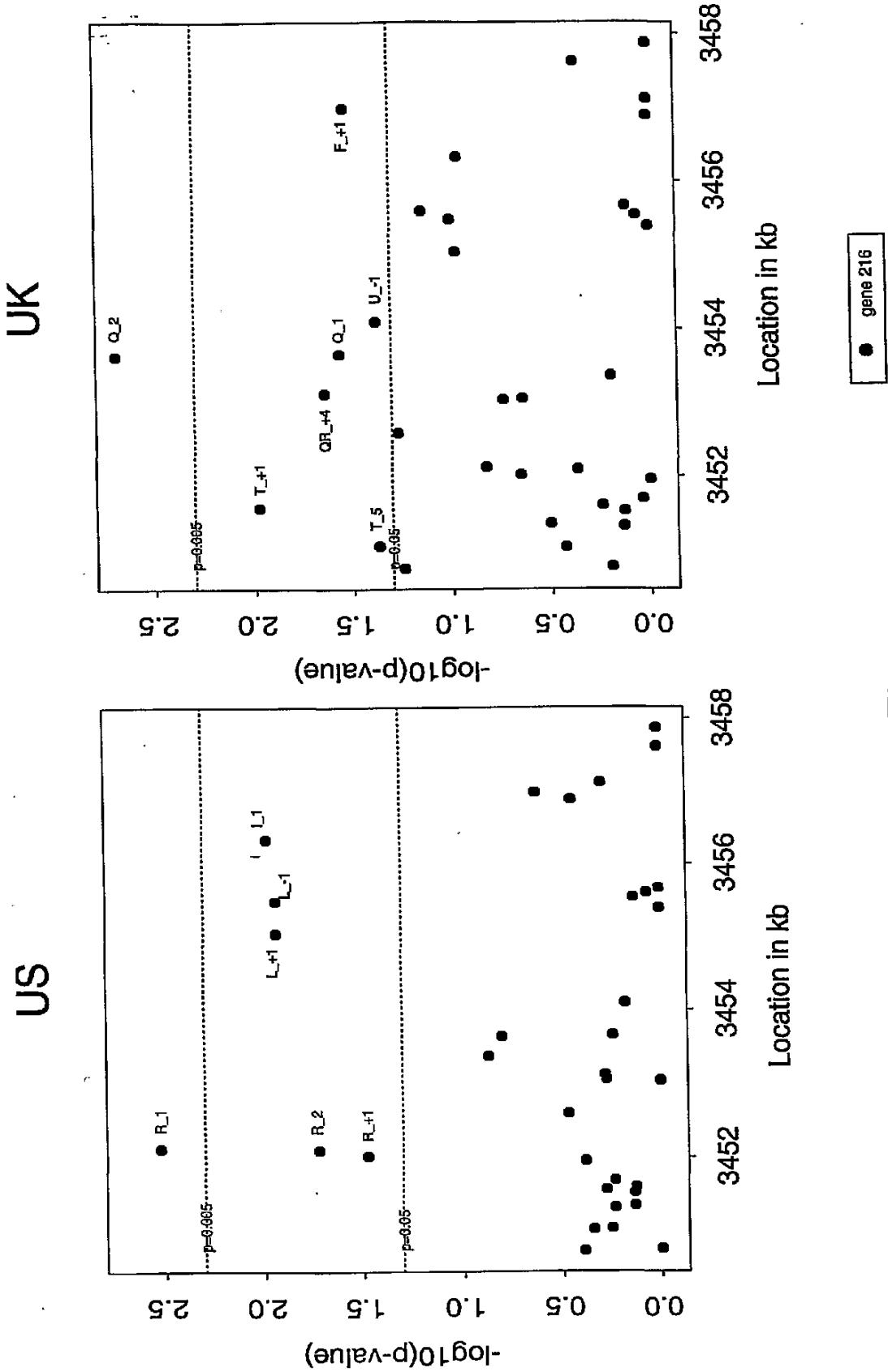


FIG. 26



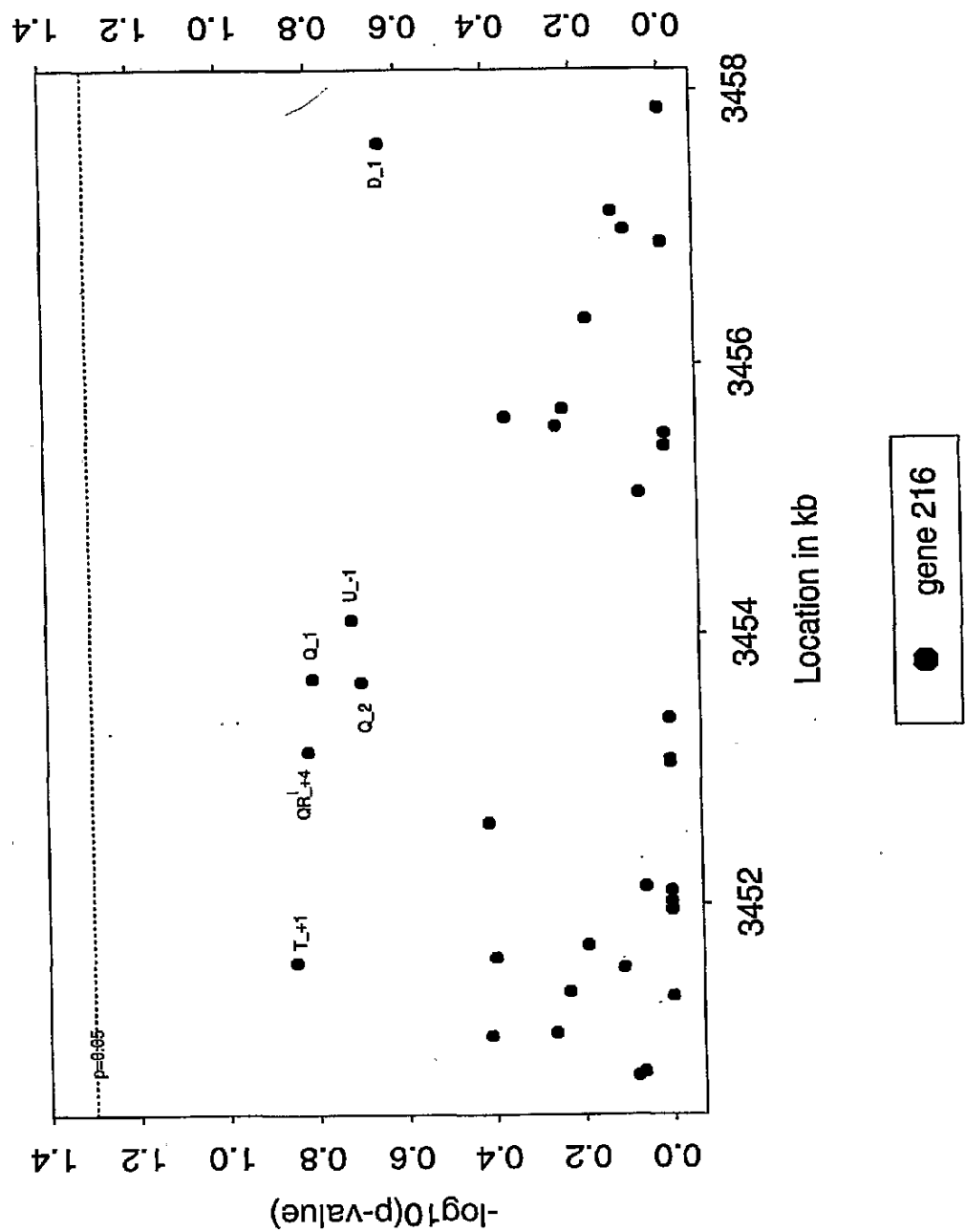


FIG. 27

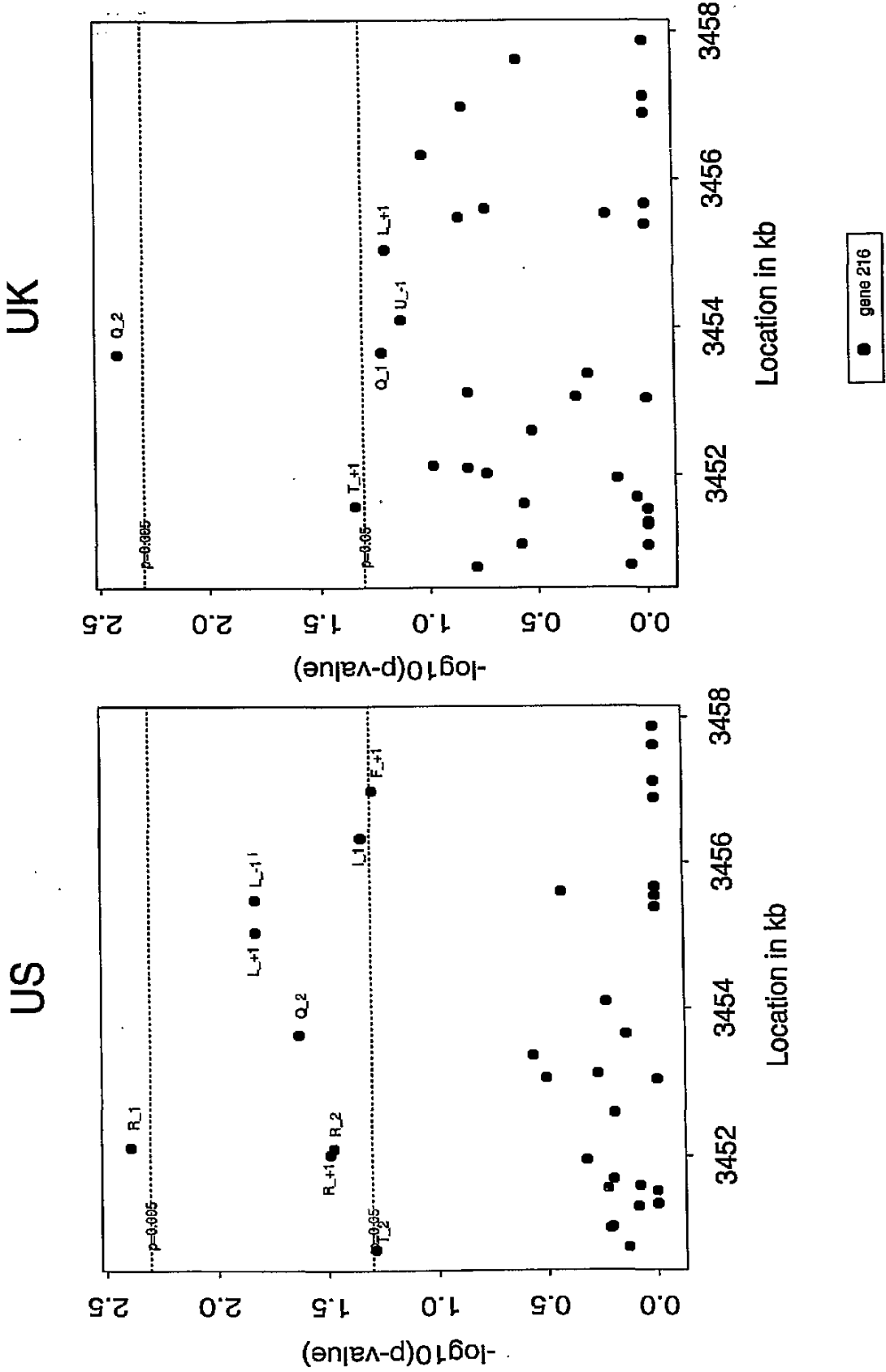


FIG. 28

10 30 50  
CGCAACTGGAGTGAGTCTCTTTCTTGCAACCCAAACAGCTGGGCTAGAACAGAGGAGTCA  
-----+-----+-----+-----+-----+-----+-----+  
70 90 110  
CATGTTAGTGGGGTTTATTCTTAGACACCATCTGTAGGAGGGGGCCAGGGGCTGCAAGTT  
-----+-----+-----+-----+-----+-----+-----+  
130 150 170  
ACATGACATTTATCACAATTACCTTACGCTGAAGTGATTGTGTCCCTTTCTGCTCCTT  
-----+-----+-----+-----+-----+-----+-----+  
190 210 230  
GAGGCTTCGTGTGTGCTCACATTTGCCACCCACAATGACCAGCCCCGTGCATGGCACGC  
-----+-----+-----+-----+-----+-----+-----+  
250 270 290  
AGTAGGTGGGCACGTGAACCTAAGCGTTTGTCTTGTCTACATTGCTGACAGTGGGGAGGGT  
-----+-----+-----+-----+-----+-----+-----+  
310 330 350  
GGAGAACATGGACTAAATCCAGAGGTGGGACAGGGAGAGAGCAAGTCAAATGCAGCCAGA  
-----+-----+-----+-----+-----+-----+-----+  
370 390 410  
TGTTTGGGGGCTGTCTGGGGAAGAGAGGGTGTAGGGAGGACCATAAGCAGCTGCTGAAGC  
-----+-----+-----+-----+-----+-----+-----+  
430 450 470  
TGCGGCTACTGGGGGGATTGTTGGGAATATCCGGCTTAGGAGTCTTCACCTGCTTCTGCATA  
-----+-----+-----+-----+-----+-----+-----+  
490 510 530  
CTTCCAAGGGGACATATGTCCATCTTGGATTGGAGGCTCTCAGAGGGGTTAGTTGTGTCC  
-----+-----+-----+-----+-----+-----+-----+  
550 570 590

FIG. 29

GCTTCCAGTGGTCTGAGGCCTCCTGAGACTTTTTGTACATTCAATTAAAGTGTAACATAC  
-----+-----+-----+-----+-----+-----+

AGAAAAGTTCACACATTATAAATGCAACTCAAGGACTTTTCCAAAAGCGAACACACCCAG  
-----+-----+-----+-----+-----+-----+-----+

ATCAAGAAATAGACCATCCTACAGTCCCCCTTACACTCTGTACCAGTTGCAGCCCCAC  
-----+-----+-----+-----+-----+-----+

AAGGGTAACTACTGTCCTTGA

TACATAAGTAGAATCACAGAGTGTGTACAATGACTTTGGAAAAGTGTGTTGACAATATCTA  
-----+-----+-----+-----+-----+-----+

TTAAAGCTAAATACCCCTTGCCCTATGAACCTGAAATTCCACCCACCTTGCCAAGGGACA  
-----+-----+-----+-----+-----+-----+

AAAAGTCCCCTCTAAATGCACCAGGCTGTCAGGGATGAAGCGTTGGCTTTGGGGCCCCC  
- - - - + - - - + - - - + - - - + - - - + - - - + - - - + - - - + - - - +

ATTACACACATGACCTTTTCTGGGGCACCCAAGCATCAGCCTGTCGTACCCAGGTGCCA  
-----+-----+-----+-----+-----+-----+

CCCTGGCGATCTCTGAAGGCTGGAGTCGGAGTGCCTCCCTCAGACATCCTGTTCTGCGTC  
-----+-----+-----+-----+-----+

ACTCC<sup>+</sup>TTGGGAGAAGTCGTGTTTACAGATGGTGGGTGTCACCCATGCCAAGCACTTCTAA<sup>+</sup>

FIG. 29

```

      1150                      1170                      1190
GGGTTAATGCTCACTGGTTTGCCTGGTTCCCAGGACATTTCTTGATGCCCTCTGGAGGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1210                      1230                      1250
TGACGCCAACAAAGCCAGTGGAGAAGCCATCTTTCCCAGGTGCTGTTCAGGCGCCCCGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1270                      1290                      1310
CTGCTCGGTGCATCCTAGGATCCCTCTTCTCCTCAGCTTTGGTTTGATGGCCTCATCTCCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1330                      1350                      1370
CCCTGCAACCTCAAATGTAAATAAACCTTTCTCAGAGACTTCGGCAGAAAAATTCCTCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1390                      1410                      1430
GACCTGCACTTGGAACAAGCTCATCTGGGTTTGGGAGGTGTCAACTGTGTAAGGATGACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1450                      1470                      1490
CTGATCCCCATGTGGCTTTTCGACTGTGTCCCCTCTACAGTCAGTTATTAGCACTGACTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1510                      1530                      1550
TGCTAGGAAGTGAGCAACACACATATCCCAGACCACATGGAGCTCAGGAGCTTGGGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1570                      1590                      1610
AGAGACAGGGAAGTGGACGACTACAGGGCCCTTCTGAAACCTGTTGCAGGGAGAAGTGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1630                      1650                      1670
GTCAGGGGATGCTAACCTGGCTTTGGGTAAAGGGACAGCCTCTGAATGACAGGACATTAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
      1690                      1710                      1730

```

FIG. 29

97/157

GCCATGGCCTGCAGTTTAAGTAGGAGTTGGCCAGTTCGAGGTAAGAATACCAGTAAGCAA  
-----+-----+-----+-----+-----+-----+-----+

1750                      1770                      1790

GAACGCCAGAGTAGCTCCTCGAGCTGCCTTCTGTACCTGACATCCACACTGAAGCCAGCC  
-----+-----+-----+-----+-----+-----+-----+

1810                      1830                      1850

CCTCTGTGTTCAGCCTTGCTTTACTGAAGAGGTGTCGCTGAGGGGCTGCTCTGGGGTGCT  
-----+-----+-----+-----+-----+-----+-----+

1870                      1890                      1910

GCTCTGCTTTCCTGTCCCCA ACTTGTTCGAGCTCGAGCCACCTCCATACTGGTGCTCCT  
-----+-----+-----+-----+-----+-----+-----+

1930                      1950                      1970

GGTTCTCAGGCCTTTGAACTCAAACTGAATCACACCACTGGCTTTCCTCGTTCTCCAGCT  
-----+-----+-----+-----+-----+-----+-----+

1990                      2010                      2030

TGCAGATGGCAGATTCGGGA ACTTTTTGGCCTCCATAATCACGTGAGCCAATTGCTATAA  
-----+-----+-----+-----+-----+-----+-----+

2050                      2070                      2090

TAAATATCTCTCTCCCTCTTTCTTCCTCTCTCTCTCTCTGTGCAAATATAGTTCCAATTA  
-----+-----+-----+-----+-----+-----+-----+

2110                      2130                      2150

TAAGAGCCCCTAACTGGA AAATAACCC TATGGTGCAC TGCTGAGTAGAGAACTGTGGTT  
-----+-----+-----+-----+-----+-----+-----+

2170                      2190                      2210

CCCTCAAAC CACCGAACACTATT CAGCAATACGA AGGAACAACTATTGATATGCAAAAT  
-----+-----+-----+-----+-----+-----+-----+

2230                      2250                      2270

AGTGTAAATGAATCTCAAAA ACATCGGAAAGAG GGAAGGAAGCCAGACACAGAAGAGTGC  
-----+-----+-----+-----+-----+-----+-----+

FIG. 29

2290 2310 2330  
ATGCCGCATGATTCCATTTATATGAAATTCTAGAACAGGCAAACTTATCTATAGACAGA  
-----+-----+-----+-----+-----+-----+-----+  
2350 2370 2390  
GAACAA CAGATCAGTGGCTGTCTGGGGTTGGGAGTGGGGAAGTTTGGCTGGAAGGGCACA  
-----+-----+-----+-----+-----+-----+-----+  
2410 2430 2450  
GGGCTCTTTCTGTGAGTGAGGGAATGTGTCTGCATTATAGTGATGCTTATGTAGTTATAT  
-----+-----+-----+-----+-----+-----+-----+  
2470 2490 2510  
ACACTTATCGAAACTCATCTTACTGGCCACTTAAAAATAAGTGCATTTTATTGTGTGTAA  
-----+-----+-----+-----+-----+-----+-----+  
2530 2550 2570  
TTATACCTTAATGAAGTTGATTTGAAAATCCAAAGTAGTAATAATAAGTAATAATCTCGT  
-----+-----+-----+-----+-----+-----+-----+  
2590 2610 2630  
AGCTGGACAGCTGTGGTGA CTCACTCCTGTAATTCCAGCGATTTGAGAAGCTGAGGCAGG  
-----+-----+-----+-----+-----+-----+-----+  
2650 2670 2690  
AGGATCACTTAAGATCAGGAGTTCCTTTTATTTTTATTTTTATTTTTTGGAGACGGAGTTT  
-----+-----+-----+-----+-----+-----+-----+  
2710 2730 2750  
CGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCGCTGCAACCTCCACC  
-----+-----+-----+-----+-----+-----+-----+  
2770 2790 2810  
TTCTGAGTTCAAGCGATTTTCCTGCCTCAGCCTCCCAAGTAGCTGGAACCTACAGGCGCTC  
-----+-----+-----+-----+-----+-----+-----+  
2830 2850 2870

FIG. 29

ACCACCA TGCCCGGCTAATTTTGTATTTT TAGTAGAGATGGGGTTTCACCATGTTGGCC  
-----+-----+-----+-----+-----+-----+-----+

2890 2910 2930  
AGACTGGTCTTGAACCTCCTGACCTCCAGTGATCTGCCCCCTCGGCCTCCCAAAGTGCTG  
-----+-----+-----+-----+-----+-----+-----+

2950                      2970                      2990

GGATTACAGGCATGAGACACTGCGCCTGGCCAAGACCAGGAGTTTGAGACCAGCCTGGGA  
-----+-----+-----+-----+-----+

3010 3030 3050  
AACAAAGTGAGACCCCTGTCTACAGAAAAATTAAAAATTTAGCTGGGCGCTGGTGCCGT  
-----+-----+-----+-----+-----+-----+

3070                      3090                      3110

GTGCCTGTAGTTCCAGCTACTCAGGAGGCTGAGGTGGGAGGATACCTTGAGCCCAGGATT  
-----+-----+-----+-----+-----+

TCAAGGCTGCAATGAGGCATGATCAGGCCACTGTCCCTCTAGCGTGGGTGACAGAGTGAGA

3190                      3210                      3230

CCCTGTCTCTAAATAATAATCATAAGAACAACAAGGACCCTCTAAACGCAC TGATATCTA  
-----+-----+-----+-----+-----+-----+

3250                      3270                      3290

. . . . .

AGGTGTATTAAGCGACCAAAAAAAAAAAGAAATCAAAGTGCAGAAAAACGTTAATAAGA

- - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +

3310 3330 3350  
GAAAAAATATGTCTGTATTGTCTTGAGTGTGAAAAATAATCTAAAAGCCTATGAAAGA  
-----+-----+-----+-----+-----+-----+

3370                      3390                      3410  
AACTAATCATATTGGTTTCCTGTTGGTGAGGAGGGCTAAGAGCACGGAGACTTTTCCCTA

FIG. 29



```

3430              3450              3470
TGCTTTCTGTACTTTTTGATTTTGAGATATGTGAATGTAGGTTTCTCTCACTGCTCGAAC
-----+-----+-----+-----+-----+-----+-----+
3490              3510              3530
TTTCACTAACCAAATTACTACATTCCAAATTCTCAAAAACAAATAGATTTACTTAAAAGT
-----+-----+-----+-----+-----+-----+-----+
3550              3570              3590
AGGCTGGGTGCGGTGTCTCACGCCTGTAATTCCAGCGCTTTGGGAGGCCGAGGCGGGCAG
-----+-----+-----+-----+-----+-----+-----+
3610              3630              3650
ATCACCTGAGGTCGGGAGTTTCGAGACCAGCCTGACCAACATGGAGAAAACCCCATCTCTAC
-----+-----+-----+-----+-----+-----+-----+
3670              3690              3710
TAAAAATACAAATTAGCCAGGCGTGTTGGCGAATGCCTGTAATGCCAGCTACTCGGGAG
-----+-----+-----+-----+-----+-----+-----+
3730              3750              3770
GCTGAGGCAGAAGAATCACTTGAATCTGGGAGGCAGAGGTTGCAGTGAGCCCAGATCATG
-----+-----+-----+-----+-----+-----+-----+
3790              3810              3830
CCATTGCACTCCAGTCTGGGTAACAAGAGAGAAACTCTGTCTCAAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+
3850              3870              3890
AAAAGATTGCTTAAAAGTTAACATCTCCGGCCGGGCGCGGTGGCTCATGCCTGTAATCC
-----+-----+-----+-----+-----+-----+-----+
3910              3930              3950
CAGCGCTTTGAGAGGCCGAGGCGGGTGGATCACGAGATCAGGAGATTGAGACCATCCTGG
-----+-----+-----+-----+-----+-----+-----+
3970              3990              4010

```

**FIG. 29**

101/157

CCAAAATGGTGAAACCTCGTCTCTGCTAAAAATACAAAAGTTAGCTGGGGGTGGTAGCGC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4030 4050 4070  
GCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCAGGGAGT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4090 4110 4130  
CGGAGGTTGCAGTGAGCCAAGATCGCGCCGCTGCACTCCAGCCTGGCGACAGAGGGAGAC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4150 4170 4190  
TCCATCTCAAAAAAAAAAAAAAAAAAAGTTAACATCTCATCCAAATTTGCACCGAGTA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4210 4230 4250  
GGAAAACAAAAGTTTAAAACATGAAACAGATGTTACTGAGGCCGAAAGGGTCTCCCAGGC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4270 4290 4310  
CTGGGAGTCTGCAGCTTTTATGCAATTCTGCCCTCTGGCCACCGCCAGGGAAGAAAGGTT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4330 4350 4370  
GTCTCCGTCTGCTGCATCGCCTTTGCCAGCAATGAAGCCCCAAGACAGCGGCAGCCGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4390 4410 4430  
TTGCCTGAACCTTCCTATCCTTGGGGGCACCCAGTGCAGGTGGATGACCCGACTCAACCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4450 4470 4490  
CCGCCAGGGCACCCTCGGGGCAGGACGGGTAGCAAGGAGGGGACAGAGATCGGCCCCAGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4510 4530 4550  
AGACCACGGAAGATCGCGCTCCTGGGGCCAACTTCAGCAGCGAGAGGCGGCCTTTGCCCA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

FIG. 29

102/157

4570 4590 4610  
CCGCCTCATCCCACCACGCCGCGGTCTCTCCAAGAACCTTCCCAGCGGTTCTCTCTCTCTCTC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4630 4650 4670  
TCAGGAGTAGAGGCCCTCTGAGACCGACGGGGAGGGACGGCTCGGGCCGGTCATCCGAGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4690 4710 4730  
GGCCGCACGGATTCCCTCCTCCGCCCAGCTCCACCCCCTCGAGGGGCGGCGGTCCGGGAG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4750 4770 4790  
TGGCGACCCGGCTCCCCCATGGCGCGCGCCGTCGGGGCCCCCTGGCCAGGCTCCGAGCGGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4810 4830 4850  
GTTGGCGGGGAGGGGAGGGCGGGAGCGAGGGCGGGCGGTGGGAGGTGGGGGCGGGAAGGTC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4870 4890 4910  
CGAAGGCGGCGGCCTGAGGCTGCACCGGGCACGGGTCCGCCGCAATCCAGCCTGGGCGGA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4930 4950 4970  
GCCGGAGTTGCGAGCCGCTGCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
4990 5010 5030  
CGGAGAGCTCGGGGGACCCCGTTGCTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
5050 5070 5090  
CCAGGCGCCGGGGTGCTTCAAGGTGAGGACGCGGGCGGGGTGCGCCCTGAGGGGCGAGGCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
5110 5130 5150

FIG. 29



5710                      5730                      5750  
CAGCTGTCCTCGGGGTAAACCCCTCCAAACAAGATGTTAAGATGGGGCTGGAACAACC  
-----+-----+-----+-----+-----+-----+-----+

5770                      5790                      5810  
TCTGCAAGCGGGTGGGAGGATTAGCCAGTCCTGCACAGCAAGTGCCCTGGCCGGGAACAGG  
-----+-----+-----+-----+-----+-----+-----+

5830                      5850                      5870  
GAGGGCAACCAGGGAGGGGGCATGCCGGGCTGGGCTGTGCTATGCAGACTGGGCGGTGGC  
-----+-----+-----+-----+-----+-----+-----+

5890                      5910                      5930  
TTCCACAGCACTGTGTGGGGACCAAACAGGTACAGGGGCCTGGTCTGTCTCTGGCCCCAGG  
-----+-----+-----+-----+-----+-----+-----+

5950                      5970                      5990  
GGAGGGCCCCAGGCGGTCCACTGCTCCCTCCCCTCTGAGCCCTATCCTGGGGTCAGGGGA  
-----+-----+-----+-----+-----+-----+-----+

6010                      6030                      6050  
GGTGATGGGACCCCTGGGAGAGGGGCGTCTATGTGCCCAATACCAGCCTGGCTCCCTCGG  
-----+-----+-----+-----+-----+-----+-----+

6070                      6090                      6110  
GTTCCACCCCCATTACCCGGTCACCGGAGCTCCAGCTCCAGCTCCAGCTCTGCCCCCTCT  
-----+-----+-----+-----+-----+-----+-----+

6130                      6150                      6170  
CTCCCTCATTGGGGTCAGGGTGCCCGTGGCCAGCACGTGCGCGCAAGGCCATGTGGACAG  
-----+-----+-----+-----+-----+-----+-----+

6190                      6210                      6230  
CACCCACACACCACACTGCACCCACACCACACCTGTGCCCGGGGCCACCCCTACCTCTTCC  
-----+-----+-----+-----+-----+-----+-----+

6250                      6270                      6290

FIG. 29

CCAAACCCTTAGAGGCCCTAGGAGCAGCAAAGCTTGGTTCTCTACTCTCAGTTAAGTGCTC  
-----+-----+-----+-----+-----+-----+

6310                      6330                      6350

TCTGGGCTGAGAGACCTCCCCTCCTTCCCCTCCCCCACATCCA CT CAGAGCCCTCCCTGC  
-----+-----+-----+-----+-----+-----+

6370                      6390                      6410

ACTGGCCCCCTCTAGCCTCCTTTCCAAGGTGGCAGACTCCTCTCGGCCCTCATCTGCCTGA  
-----+-----+-----+-----+-----+-----+

6430                      6450                      6470

TGGCAATTCACTCATCCAATCAAGGAGGGCTTCTTGGAGGAAGGGTCTTTGATGTTTGTA  
-----+-----+-----+-----+-----+-----+

6490                      6510                      6530

GTCTGGGAGAGAAGGTGGAGGAGAAAAAAGGAGTTGGGGTGGCCTAGCAGGAGCTGAGTC  
-----+-----+-----+-----+-----+-----+

6550                      6570                      6590

ACTTCCACAGGCAGCCATCAGCCCAGCAGGACTGAGGCCAGGGCTGCGTGGAGGGGGGAG  
-----+-----+-----+-----+-----+-----+

6610                      6630                      6650

GCTGTCTGTTCTGGGAGCTGGGACTGGGTACCGGGGGAAGGAGGGCTGCTGCAGGCTCTG  
-----+-----+-----+-----+-----+-----+

6670                      6690                      6710

GGTGCCTGGGGCCTGGCTCCTGCAGGGCGGGCCTGTGAGAGTGGTTGGGGCCAGTGGAGG  
-----+-----+-----+-----+-----+-----+

6730                      6750                      6770

GGCTGGGAGCATTCCAGGGGAACATTCCAGGCGCCCTCTGAGTAATGCTTGGCTCTGGGA  
-----+-----+-----+-----+-----+-----+

6790                      6810                      6830

TTCTCTCTAGAGCCCCCTTAGGCACACCCGGCCAGGGAGCACCAAGGCTCCGTCCGGAAG  
-----+-----+-----+-----+-----+-----+

FIG. 29

106/157

```
      6850              6870              6890
      .               .               .
CGTCCCCTCCCCTTGAAGAGATGAGGAGGGGCCTTCTGGGCCAGGGTACCAAAACCCTGC
-----+-----+-----+-----+-----+-----+-----+
      6910              6930              6950
      .               .               .
CACCAGGACAGAGTCCCCGAGGGAGCTCTGGGCAAGGTGGACCTCGCAAGGCAACATCTG
-----+-----+-----+-----+-----+-----+-----+
      6970              6990              7010
      .               .               .
GCTGTTGTTTTTCTCAGATGATGGGGGGGGCACAAGTGTCTCTCTTCGTACATCTCTCA
-----+-----+-----+-----+-----+-----+-----+
      7030              7050              7070
      .               .               .
CCCTAAAGGCATCTGCTGCCCATCTAAAAATCCCTAAGGCTGCCGCGCTCTTTCCTTCCC
-----+-----+-----+-----+-----+-----+-----+
      7090              7110              7130
      .               .               .
CTCTGCACTGGCGGCCTTGGCCTCTTCCTTGTGATCGCCGAGCCCAAGCCTGCCCCCGA
-----+-----+-----+-----+-----+-----+-----+
      7150              7170              7190
      .               .               .
CAAAGGTCAGGGGACTCCCGTGTCCCCAGCTGAGCTGTCCCTTTCAGCCTTCTCTTTTC
-----+-----+-----+-----+-----+-----+-----+
      7210              7230              7250
      .               .               .
CTCCTCCTTGATAGCTCCTCAGATCCAAGGATGCCACGGGCGTCCCTCCTTCTCCAGGC
-----+-----+-----+-----+-----+-----+-----+
      7270              7290              7310
      .               .               .
TGAGCCACGCGTGTGAAGGTGAAGTCTGCCCCAAAAGGCCTCCAGTGCCTCCCTGGGG
-----+-----+-----+-----+-----+-----+-----+
      7330              7350              7370
      .               .               .
ATGTCCTCTACCCCCCTCCCTCTGCTTTGTCCCATGCCCTGTGTTCTCAGGTCCCCCT
-----+-----+-----+-----+-----+-----+-----+
      7390              7410              7430
```

FIG. 29

CACCCCTGTGCTCTGTCTTTACTCCAGGACATATCCCTGGGCAGCCAGTCAACCCGCACTG  
-----+-----+-----+-----+-----+-----+-----+-----+  
7450 7470 7490  
GGTCCTGGATGGACAACCCTGGCGCACCGTCAGCCTGGAGGAGCCGGTCAGTGCCATGTC  
-----+-----+-----+-----+-----+-----+-----+-----+  
7510 7530 7550  
TCCCCGCCCTCCACAGGGGCCCTGAACCTCCAGCCCTTTTGTCTCTCCCTACATTACAG  
-----+-----+-----+-----+-----+-----+-----+-----+  
7570 7590 7610  
CTTCTAGTTTTGTCTGGGGTCCCCAGAACCACCAAGTCACTACTCCTATAGGCCCCCTGCCT  
-----+-----+-----+-----+-----+-----+-----+-----+  
7630 7650 7670  
CCCCTGCCCCCTCAAGTGGGCAGAAGAAGGCACTGGGGTTTGGACATCTGGATCTCGTGAG  
-----+-----+-----+-----+-----+-----+-----+-----+  
7690 7710 7730  
CCCGCACACATGGAAGTCATTTAGCTTCTCCACCCCACTCCCTCTTTCCCTCCCTCC  
-----+-----+-----+-----+-----+-----+-----+-----+  
7750 7770 7790  
CTGGATGATCTGGGGCCACCCCCACCCCCACCAGGCAGAAATGGGTCCAGAGTTTGTGGGT  
-----+-----+-----+-----+-----+-----+-----+-----+  
7810 7830 7850  
CCTGAAGCTTTTCAGGAGCCTCTAAAAAAAAAAAAAAAAAAAAAGCACCAAAAGAAAA  
-----+-----+-----+-----+-----+-----+-----+-----+  
7870 7890 7910  
CCTTTTGCAAAGTTGACCAGAACATGTGACCCTGTGGACACACTGCTGTCCCTCTCAGGG  
-----+-----+-----+-----+-----+-----+-----+-----+  
7930 7950 7970  
CCCTGCCACGAAGGCCTGAACCTTCAGCCTCACTGGCTCCTGTGGAATCCACTTCTGGTA  
-----+-----+-----+-----+-----+-----+-----+-----+

FIG. 29



108/157

```
      7990              8010              8030
TGGGGGGGGGCAGTGGTCACCTCTCCTGATGTCCCCAGATGTAAGACCACCCCATGTGCTT
-----+-----+-----+-----+-----+-----+-----+

      8050              8070              8090
CTTCTGCAGGACGCTCTGCCCCAGCCTCTTCCCAATCCCGCTCTTCACACGCTTCCAGAA
-----+-----+-----+-----+-----+-----+-----+

      8110              8130              8150
TAACCATGCCCCATCTGTTTGTGCCATAATATCTGTGCTGCAAATAAGAGGGCAGTAGC
-----+-----+-----+-----+-----+-----+-----+

      8170              8190              8210
CTTGATATGCTCATTTTACAGAGGGGCAAACGGAAGCCCAGAGAGCTTGGGGAAATTGTC
-----+-----+-----+-----+-----+-----+-----+

      8230              8250              8270
CATGGTCACACAGCTCTTTAGGCTGGGAGCCTGAGACCCACTAAGGTCTGAACGATTTTA
-----+-----+-----+-----+-----+-----+-----+

      8290              8310              8330
AACCATTGGCTACACCCCTGCCCCCTCCTAGAGAGCCCTCTTGTTTGGAAATTTTCAGCCC
-----+-----+-----+-----+-----+-----+-----+

      8350              8370              8390
TACTGTCCAAATCCAGCAAGAGGGAAGGCAGGGGAGCATTGCCATGAAGGCTGAGAGGCC
-----+-----+-----+-----+-----+-----+-----+

      8410              8430              8450
CCCAGAGACCCAGCAGCTCCCAACCCAGGGCCCTCACTGGGATCCCCTAGGCCCATAGG
-----+-----+-----+-----+-----+-----+-----+

      8470              8490              8510
CCCCCATTCCTACTGGTCAAGCACGGCACTGGCCTGAGCTTTGAGATTGCCCTCCCCATCC
-----+-----+-----+-----+-----+-----+-----+

      8530              8550              8570
```

FIG. 29

109/157

CCAGGAGGGGAAGGCTGGACACACACTGGGGTCACTCTGCCTCTGGGCCCTCCCTGTCTGT  
-----+-----+-----+-----+-----+  
8590 8610 8630  
CTGGCCTGGGCTGTGACCAAGAGGAGCCCCAAAGGGGCTCTGCTTCCCCACCGGTGG  
-----+-----+-----+-----+-----+  
8650 8670 8690  
GCCCCTGCCCCCAGGAAGCCTGCCAAGATGGTACAGAAGAAAGAGTAGAGGCTAGGTATC  
-----+-----+-----+-----+-----+  
8710 8730 8750  
CCCTCCAAAAGGCAGGAAACACTCACATTTCAGATGAGGGGTATATATCAAGGGGCAGG  
-----+-----+-----+-----+-----+  
8770 8790 8810  
GTACCAGGAGGGCAAGAGTAAGATAGCAGGGGCTGCAGAGGAACAGGGACCTCGAGTAT  
-----+-----+-----+-----+-----+  
8830 8850 8870  
GGCCTTTTTCCCGGTGCAGACCTTTCCCCAATAAAGCAAGTGGCATTCAGCCTCATGAG  
-----+-----+-----+-----+-----+  
8890 8910 8930  
CTCATGCTGGAGGCCTTGTGGGGCCTGTGGCCAGGGAGGCAAGGACCATCTGCTCCCCAC  
-----+-----+-----+-----+-----+  
8950 8970 8990  
TTGCGAAGGAAGAACTCCCTCCAAAGACTCTGAGACCCTTGGACAGGGCCCCAGGCCAGT  
-----+-----+-----+-----+-----+  
9010 9030 9050  
GCATTTTTTGAGAAAAGGAGTCGGGGGTTAAACATTCCGAAGGCGCAGCAGCCTCCCAGG  
-----+-----+-----+-----+-----+  
9070 9090 9110  
AAGCTCCTGGGCGGCTCCAACCTCTGGGCCCCCAGCCAGGCTGAGTGGACAAGGGGGAAG

9130 9150 9170  
TGGGGTGTTCCCAAGGGTGGGAGACGCCAAGAGGGTGGGGGAAGGAGAGAGGGCTGGCC  
-----+-----+-----+-----+-----+-----+-----+  
  
9190 9210 9230  
GTCCAAGCCAGCCTCCTGACACCTAGCTGAGAGCCAGTGTGCTCTCTTGGCTGGAATGGC  
-----+-----+-----+-----+-----+-----+-----+  
  
9250 9270 9290  
GTCCATGTTTTACTTCGTGGGTCCAGTGAAGCAGGTGTCTCGGAGCCGGAGGGACGGGGGGCTG  
-----+-----+-----+-----+-----+-----+-----+  
  
9310 9330 9350  
CTGGAGGCCCCAGGAAAACCTTTGGAAGAGGGAGCAGTTTGCCAAATTTGGAAGTGGAGGAG  
-----+-----+-----+-----+-----+-----+-----+  
  
9370 9390 9410  
TCAAATTTGAATTCTATAGGAAATGAGCAGCAGCTCATTTGGAACCAAGCCTCAGGTAGC  
-----+-----+-----+-----+-----+-----+-----+  
  
9430 9450 9470  
AGAGGCTCTGAGGAGGCCCTGACCATGGCTACCCGATGCCCCATAATGTCCTCAGCACC  
-----+-----+-----+-----+-----+-----+-----+  
  
9490 9510 9530  
CCTCTGTCTTCCCCTGCTTTTGATGCCCCCTTCTGGGCATGAAAGAAGAGGGCGGGGCCAG  
-----+-----+-----+-----+-----+-----+-----+  
  
9550 9570 9590  
GGGAGGGGCACCTTTCTGGGACCTCTGGTCTCTAGGGAGGATGCTGGTGTGCCTGGCAGG  
-----+-----+-----+-----+-----+-----+-----+  
  
9610 9630 9650  
CTGTGCCAACGCCCTTCCAAGTGGCTGTTGTTCAGGACTGCAAACATCCTGAGTTTGGGAA  
-----+-----+-----+-----+-----+-----+-----+  
  
9670 9690 9710

FIG. 29

CATCTTTGTATGTTCTCACCTCCTCCACGCCCTCCATAGTATGTGGGGGGTCCTGCTGAC  
-----+-----+-----+-----+-----+-----+-----+  
  
9730                      9750                      9770  
  
TCCCCAGCCCACGTTCTCCCAAGAACTTCCTCCCCAGCCGGCTCCACAGGCCACCTACT  
-----+-----+-----+-----+-----+-----+-----+  
  
9790                      9810                      9830  
  
CCCTGGCAGGCAGGAGGCCTGGAGGCCACCATCTCAGCTCCACACTCTTTCTTGCCAGG  
-----+-----+-----+-----+-----+-----+-----+  
  
9850                      9870                      9890  
  
TCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAGGCCAGGAGCTCCTGCTTG  
-----+-----+-----+-----+-----+-----+-----+  
  
9910                      9930                      9950  
  
AGCTGGAGAAGAACCAGTGAGTGCCAGGCTGGGGTAGGGCTGGGAGGAGGGGATCAGTGT  
-----+-----+-----+-----+-----+-----+-----+  
  
9970                      9990                      10010  
  
TGGGGGGCAGGGACTGACACAGATCTGTGCGGGTGGCTGGATGGGCAGAGGACCCCAGAG  
-----+-----+-----+-----+-----+-----+-----+  
  
10030                      10050                      10070  
  
AGGGTGCAGATGACAGGGGAGAGTCACGCAGGCCTGTGGTTGGCTCCCTGGAGGCTGAAGA  
-----+-----+-----+-----+-----+-----+-----+  
  
10090                      10110                      10130  
  
GGACCGCTGAGGCTGTCAGCCCCGCTGTGGGGCACCTCCGCCCTCCCAACCCCAGGAGCG  
-----+-----+-----+-----+-----+-----+-----+  
  
10150                      10170                      10190  
  
GCTTGTTAGCTCCCTGCTGGCGATGAGTGAGCACCACCTAGTGGACATTGTGAAGATATG  
-----+-----+-----+-----+-----+-----+-----+  
  
10210                      10230                      10250  
  
CTGAGTCTAAAGAAATCCTAGAGGGAAAAGATGAGCCGGCACCCCAGGCTAAGGGAATGG

FIG. 29

112/157

```
10270      10290      10310
CAGGGACCAAGATGCGGTGGCTTTGGGAGGCCGAGGCGGGCGGCTCACCTGAGGTCAGGA
-----+-----+-----+-----+-----+-----+-----+

10330      10350      10370
GTTTGAGACCAGCCTTGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATT
-----+-----+-----+-----+-----+-----+-----+

10390      10410      10430
AGCCAGGCGTGGTGGCGGCGCCTGTAATCCAGCTACTTAGGGGGCTGAGACGGGAGAAT
-----+-----+-----+-----+-----+-----+-----+

10450      10470      10490
CGCTTGAACCCCGGAGGTAGAGGTTGTGGTGAGCCAAGATCACACCACTGCACCACTCCG
-----+-----+-----+-----+-----+-----+-----+

10510      10530      10550
GCCTGGGCAAAGAGTGAGACTCCGTCTCAAAAAAGAGAAAAAAGAAAAAGAAAAAA
-----+-----+-----+-----+-----+-----+-----+

10570      10590      10610
AAAAGAAAGAAAAGAAAAAGAAAAGATGCAGTGGCTACACTTGGGGGCAGCAGTTTGT
-----+-----+-----+-----+-----+-----+-----+

10630      10650      10670
CTGACCTGCCTGGAAGGTCTCCATCTACAGGGAGGGGAGCAGGGGGGAATGAATTTGGAG
-----+-----+-----+-----+-----+-----+-----+

10690      10710      10730
AGTCCCAGGAGGGGCCAGATCACAGAAGGCCATTTTGGTGCTCAGTGTCTGACCATCCA
-----+-----+-----+-----+-----+-----+-----+

10750      10770      10790
GAGCCAAAGATTTTGAGCTGGGGAAGGGACAGGCAGACCTGTGCTCAGGAAGGTGCCTTG
-----+-----+-----+-----+-----+-----+-----+

10810      10830      10850
```

FIG. 29

GGCTGGGTGGGGTGGGTGTCCGGGCTGGAGCGCAGGCTCTTAAACCACCCAGATTATGT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+

10870 10890 10910  
TATCAGTATATATCACCTACTGAGTGCTTGACCGCAGGCGCTGTTCTGAGCACCTTGACAC  
-----+-----+-----+-----+-----+-----+-----+

10930                      10950                      10970

GTATTTTATTCTCCCTCGTGAGTCGGATGGACAGGGAACAACCTCTAGTTCCACTGTGC  
-----+-----+-----+-----+-----+

10990 11010 11030  
CCAACCATATTTTCCCGACGTCCCTACCCCTTCAATGGGGTGGTCACATCACCTACCTCC  
-----+-----+-----+-----+-----+

11050                      11070                      11090  
 .                      .                      .                      .  
**TAGGGTGGCGGGTGTTGTGTTGGGCAGGGCTAGGGGCAGAGCTGGGGCAGGTGGTGGAAT**  
 -----+-----+-----+-----+-----+-----+-----+-----+-----+

11110                      11130                      11150

GCCTGGGAGGGGGGAAGCAGCCATCATTAGCGGGTGGTCTGGAGGTAATGAGGCCAAGGT  
-----+-----+-----+-----+-----+-----+-----+

11170                      11190                      11210  
       .                      .                      .                      .  
**GAGGT TGGGTTAAGGATTTCTTTAAAGAAGACAGATTGACTTATGATTGATCCATCCGT**  
 -----+-----+-----+-----+-----+-----+-----+-----+-----+

11230                      11250                      11270  
GTGGGAAAGATCCTGTTGAGATGGAGCCTGAAGATGGAATCATTACCGGAGTGGGTGTGG  
-----+-----+-----+-----+-----+-----+-----+

11290                      11310                      11330  
AGAAAGGCAGGGAGGGTGAAGCAGCGTGCGCAGGTGGCGATTCTGTTTTCTCTGGAGGCA

11350                      11370                      11390  
GGGGGTGAGCATCAATCACTGAAGGACAGGTGGGAGGTATGTGGGGTCTAGAAGTCTGAG  
-----+-----+-----+-----+-----+-----+-----+

FIG. 29

114/157

|   |       |       |
|---|-------|-------|
| 11410   | 11430 | 11450 |
| GAAAATATTTCAAGGATCTAGGGCAGGTGGGGGCAAGAGGGTCGACCAGATGCCCAACAA                              |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11470   | 11490 | 11510 |
| AGGAGGGCAGCAGGCAGGGGAACTGGGGGAGGTACCCGCATTTCCCCAAC <sup>1</sup> TCCAAGTCCC                |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11530   | 11550 | 11570 |
| ATTCTTCGGCAGTGTCTCCTGACTCCTCCCCCTCCCGATCCTGTGGATCCTGCTGCCTGCT                             |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11590   | 11610 | 11630 |
| GCAGGTCCCCTGGGAACCA <sup>2</sup> CAAACTCTTCCCCTATTCCCAC <sup>3</sup> TCCTCCCCGGCGTCCCTCCC |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11650   | 11670 | 11690 |
| TGGTGCTTCCCATATTCA <sup>4</sup> CACTCTCCCACTAAGCCATCACCAAGGCTCCTTCCTCTAG                  |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11710   | 11730 | 11750 |
| CCCCAAGAGTTTCTGATCTGAGCAAGTCACCATTGCTCCTGTCCCTTCCCTAAGACACAC                              |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11770   | 11790 | 11810 |
| TGTGAGTGTCTCACTCATAAAGCTGCTCCATTAGCATTTAGGGAGGAAGGCTGGGAGACA                              |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11830   | 11850 | 11870 |
| TCCTGGAGGAGGCAGGAGGAAGCTGAATTCAGTGTTCCTGTAAACACCCCTCTCAGCAG                               |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11890   | 11910 | 11930 |
| GCTGCTGGCCCCAGGATACATAGAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGTCT                             |       |       |
| -----+-----+-----+-----+-----+-----+-----+  |       |       |
| 11950   | 11970 | 11990 |

FIG. 29





116/157

12550 12570 12590  
AGACAGGAATTACAGGCCTCGAGTCCCAGTATTTTTATTGAAGTCTGAAGAAACAAGTT  
-----+-----+-----+-----+-----+-----+-----+  
12610 12630 12650  
CCAGAAAACATGTTAAACTTCCTTCTGGGAGCTGGGGTTGGGGGTCAAGCCCA  
-----+-----+-----+-----+-----+-----+-----+  
12670 12690 12710  
GCAGCTTCCACTCAGGGTCCCCATTGTGCACCTCCGCAGGGCAGGCGAGAAGCGCGCAGGA  
-----+-----+-----+-----+-----+-----+-----+  
12730 12750 12770  
CCCGGAAGTACCTGGAACGTGTACATTGTGGCAGACCACACCCTGGTGAGGAGAGACCCCA  
-----+-----+-----+-----+-----+-----+-----+  
12790 12810 12830  
GGGGTTGGCGGGGTCAGGGATGGGGCCAGCTCAGCCCCTCAAGCCACCGGGATTCTGCCC  
-----+-----+-----+-----+-----+-----+-----+  
12850 12870 12890  
TTCCCAGTTCTTGACTCGGCACCGAAACTTGAACCACACCAAACAGCGTCTCCTGGAAGT  
-----+-----+-----+-----+-----+-----+-----+  
12910 12930 12950  
CGCCAACTACGTGGACCAGGTTGGGGGCGGCGGGGAGAGAGCGGTGATGGGGGTGGCGGC  
-----+-----+-----+-----+-----+-----+-----+  
12970 12990 13010  
GGCAGGACAGGCAGGTGCTGGTGGGGTTTGGGGAAGAGGAAGGGCGCCCCACGAAGGACC  
-----+-----+-----+-----+-----+-----+-----+  
13030 13050 13070  
ACCGGCGCGATGGGGCGCCCTGTCCCGGCTTCAGCCCCGCCTCGCCCTCAGCTTCTCAGG  
-----+-----+-----+-----+-----+-----+-----+  
13090 13110 13130

FIG. 29



118/157

```

13690      13710      13730
      .      .      .
TCCTCCGCTGCGTTTCTTTGGTCGTCCTCAGTTTCCTCTTCTGTAAATGGGGATAATG
-----+-----+-----+-----+-----+-----+-----+
13750      13770      13790
      .      .      .
ATCATAGTGTCCGCTTCAGGGTGGTTTATGAGGCTTAAAGGGAAGAAGCTCAGGCAAAGT
-----+-----+-----+-----+-----+-----+-----+
13810      13830      13850
      .      .      .
GGATTCTCAACGGTATGAAGATTATTTTCCGAGTAACCTGGCGAGGTTACTCCTACACCG
-----+-----+-----+-----+-----+-----+-----+
13870      13890      13910
      .      .      .
GGAGGAGCACCGTCGGGTGCGGATTCCACCTTGGGTCCCCGGGCTGCTCACTATTGGGGCC
-----+-----+-----+-----+-----+-----+-----+
13930      13950      13970
      .      .      .
GCATCGTCCCCTGTCCCGCTTGTTGTGTGACTTTGCGCGGGTTACTTCCCCTCTCTGGGC
-----+-----+-----+-----+-----+-----+-----+
13990      14010      14030
      .      .      .
TCTGCGCGTCTGGCGGCTGTAGCCAAGCCCAGGGGTGGGGATCAGAGAAGCGCGGGGGTT
-----+-----+-----+-----+-----+-----+-----+
14050      14070      14090
      .      .      .
GGGGGACTGTCCCTCCATGCCCAATGCCCTCCCCGTGCCGGTAGGCACCCGTTTCCGCGC
-----+-----+-----+-----+-----+-----+-----+
14110      14130      14150
      .      .      .
GTGTTTCAGCGCCTGCAGCCGCCGCCAGCTGCGCGCCTTCTTCCGCAAGGGGGGCGGCGCT
-----+-----+-----+-----+-----+-----+-----+
14170      14190      14210
      .      .      .
TGCTTCTCCAATGCCCCGGACCCCGGACTCCCGGTGCCGCCGGCGCTCTGCGGGAACGGC
-----+-----+-----+-----+-----+-----+-----+
14230      14250      14270
      .      .      .

```

**FIG. 29**

TTCGTTGGGAAGCGGGCGAGGAGTGTGACTGCGGGCCCTGGGCCAGGTTAAGTCGGGCTCGCCCC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14290                      14310                      14330  
GCCCCCACCTTGCCCTCTCCGCTCAGGTCTGGGGCGCTGCGCCCTCACCTGGGCCCTTCTT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14350                      14370                      14390  
GCCTTTCTGGTCCCAGGAGTGCCGCGACCTCTGCTGCTTTGCTCACAACCTGCTCGCTGCG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14410                      14430                      14450  
CCCCGGGGGCCAGTGCGCCACGGGGACTGCTGCGTGCGCTGCCTGGTGAGGGCATGGAA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14470                      14490                      14510  
GGTTCAGGGTGAGGGTTTCGGGGAGCTTGGGAGCCGGCCTGTTGGCCTTAGTTAATTGGT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14530                      14550                      14570  
GCCCTCAGGTTCCCCCGTTGGGTGCTGGGCTTGGGTAGGCCTGGCTCCCCCAGCTCCGAG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14590                      14610                      14630  
CCGCGCTCTCGGCATGGACCTCTCACTGCACGTGGCCTCTCTCTGCCTTCCCCACCACCC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14650                      14670                      14690  
GTCACCTGCGCAGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCATGGGTGACTGTGA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14710                      14730                      14750  
CCTCCCTGAGTTTTGCACGGGCACCTCCTCCCACTGTCCCCCAGACGTTTACCTACTGGA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
14770                      14790                      14810  
CGGCTCACCTGTGCCAGGGGCAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGA

FIG. 29

14830 14850 14870  
GCAGCAGTGCCAGCAGCTCTGGGGGCCTGGTGAGAGGACACGAGCACCCCTTGACCCCTGC  
-----+-----+-----+-----+-----+-----+-----+  
14890 14910 14930  
CCCCCATCCTCTGGTGGGGCCAGTTTTCTACTGTGGGGAAGATGGGCAGGGGAAACTGAG  
-----+-----+-----+-----+-----+-----+-----+  
14950 14970 14990  
GCCCCTGAGCGCAGCCCCCTCTCCGAGCTGCCCCCAGCCTGGCCCATGCTTCCTCAGGCT  
-----+-----+-----+-----+-----+-----+-----+  
15010 15030 15050  
CCCACCCAGCTCCCGAGGCCTGTTTCCAGGTGGTGAAGTCTGCGGGAGATGCTCATGGAA  
-----+-----+-----+-----+-----+-----+-----+  
15070 15090 15110  
ACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGTAGGGAGTGGAGC  
-----+-----+-----+-----+-----+-----+-----+  
15130 15150 15170  
TGAGTGGAGGGAGCAGAAGCTATGGAGTGGGTTTGGGGAAGGGGGGTACTGCAGCTGTTG  
-----+-----+-----+-----+-----+-----+-----+  
15190 15210 15230  
ACCCCCCTCTACTTCCTCCCCAGGGATGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAA  
-----+-----+-----+-----+-----+-----+-----+  
15250 15270 15290  
AAGCCCAGCCTGCTCGCACCGCACATGGTGCCAGTGGACTCTACCGTTCACCTAGATGGC  
-----+-----+-----+-----+-----+-----+-----+  
15310 15330 15350  
CAGGAAGTGACTTGTCGGGGAGCCTTGGCACTCCCCAGTGCCAGCTGGACCTGCTTGGC  
-----+-----+-----+-----+-----+-----+-----+  
15370 15390 15410

FIG. 29

CTGGGCCTGGTAGAGCCAGGCCACCCAGTGTGGACCTAGAATGGTGAGCTCTGCCCCACC

-----+-----+-----+-----+-----+-----+

15430                      15450                      15470

ACCCTCCTTGCCGTTTGAATCCCGCAGGCCAGTGTCCCCCTCACTGCCTGGTGCACTGC

-----+-----+-----+-----+-----+-----+

15490                      15510                      15530

CCGTAGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCTTCAGGAGCTTCAGCGCTGC

-----+-----+-----+-----+-----+-----+

15550                      15570                      15590

CTGACTGCCTGCCACAGCCACGGGGTGAGAGCCCAGGAGTGGGGGTGACCTTGGGGTTC

-----+-----+-----+-----+-----+-----+

15610                      15630                      15650

CTAATCCTACGTGACCCTCCTCTTCTCTCTGTCAGGTTTGCAATAGCAACCATAACTG

-----+-----+-----+-----+-----+-----+

15670                      15690                      15710

CCACTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCAT

-----+-----+-----+-----+-----+-----+

15730                      15750                      15770

GGACAGTGGCCCTGTGCAGGCTGAAAGTATGCCAGTGGGGGGGCATGTGGGCAGGAGCTGG

-----+-----+-----+-----+-----+-----+

15790                      15810                      15830

GGTGGTGCACTGCTCAGGACTCAGCGCCCCCTCCCCCAATCCCCGCAGACCATGACACC

-----+-----+-----+-----+-----+-----+

15850                      15870                      15890

TTCCTGCTG GCCATGCTCCTCAGCGTCTCTGCTGCCTCTGCTCCCAGGGGCCGCCTGGCC

-----+-----+-----+-----+-----+-----+

15910                      15930                      15950

TGGTGTGCTACCGACTCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCTGCAGAAGG

-----+-----+-----+-----+-----+-----+

FIG. 29

122/157

15970 15990 16010  
GACCCTGCGTGCAGTGGGTAGGCTCCGAGCGCCTGCTTCCTGAGCCTACTCCTGCGGTTCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16030 16050 16070  
CCCTCCTCAGAGCTCTGCTGGGGCTGTGGGAGCTGGGGCAGGCCCTCAGCCTTGCCCCCA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16090 16110 16130  
GGTGCAGAGAGCAGCCCCAGAGGCCATGGAAAGAAGTAGCTTTGAACAGGAGGTTCAGT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16150 16170 16190  
GGCCTCCAGTCAAGCGAGGGGGTGGATCCCTGCCCCACCACAGCACCGCAAGGCATGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16210 16230 16250  
CCCTCTACCTCCCAGTACAGCTCCTCTTGTCCTCTCTCTGCTTCTCCCACCAGCTGGCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16270 16290 16310  
GCCTCACCTTGACTTCGCCCTGTTTTCCCTGGCTCAGATTGCAGTCCCTGTACCATGC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16330 16350 16370  
TGCCCCCGGAGGCCTGTCCAGCCTCTGTCTCACCAGTTTTCGGGCCCTTTGCCACTTCCTC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16390 16410 16430  
TGCACAAATCACCTCTGTCACCCCTTGAAGTTCCCAAATGCTGGGCCCAGCACATCTTT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16450 16470 16490  
TCACTCCATACCACTGGTCAGCTGCGGTGCTGGCTGCCCCCTGTGCCAGGGCCCTGCCTTA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
16510 16530 16550

FIG. 29

123/157

ACCCAGTTCTCTGTGACCTGGGTGGTGGCGGAGTGGGGAGTCACATAATACTAAGCATGG  
-----+-----+-----+-----+-----+-----+-----+

16570 16590 16610

CTGTCCTAGGACTCACCCTGCACCAGGGCCCTAGGCAGGGCAGGCACCTCTGTGGCCATGT

-----+-----+-----+-----+-----+-----+

16630                      16650                      16670

CTGACATAGCCTGGTCTTG<sup>.</sup>GGGAGTGCTCCGGGCAAGCCAAGGGAGATGGCATGATTGGG<sup>.</sup>

- - - + - - - + - - - + - - - + - - - + - - - + - - - + - - - + - - - + - - - +

16690 16710 16730  
CCAGAGATGGGGGCAGAGGGCATAACAGACAGGGGCAGGGCACCACCTGGGCCCCGGGTG  
-----+-----+-----+-----+-----+-----+-----+

GCAGCTAAGAGGACCCTGACAAAGCGAGTTGTGATTGAGGGTCTGTGGGCAGAGGAGCAA

16810                      16830                      16850

GGTGGCCAGAGCCTGGCGTGTCAGCACGGAGGGGCGCTGCAGAGGGTGGCGGCTGCTTC  
-----+-----+-----+-----+-----+-----+

16870 16890 16910  
TCATCCCCAGGCGGGAGTCTCAGGGCAGGGGAGAATGTTTTGAAGGAACATCACAGGAAA  
-----+-----+-----+-----+-----+-----+-----+

TGACAAGGCCTTGGGGATGGGATGGGGACAGTCAAAGATGGCTTGGAATCATCAAGGGC

16990 17010 17030  
AGCAGGGCA<sup>.</sup>CCCAGGGGCA<sup>.</sup>AAGGAGAGCAGACATAGCTGCCGAAGGGGCGGACATCCAAGG<sup>.</sup>  
-----+-----+-----+-----+-----+

17050 17070 17090  
TTCTTTGGAAGCTGAGCGATGCCAGCATCTGGAGAGTGCCAGGCTGCTGGGTGGTGTCTAG

FIG. 29



```

17110      17130      17150
AGCCTGGAGGAAATGTTAGGACTAGAGAGAGGAGGTGCCAGCCGAGGCATGAGGCTCAC
-----+-----+-----+-----+-----+-----+-----+
17170      17190      17210
TTGGAGCCTGGATCCCAAGGCTCCCCTGAAGAGGGAGCAGGAAGGGAGCTGAGAGGGTGA
-----+-----+-----+-----+-----+-----+-----+
17230      17250      17270
CTTGAGCAGATGGGTGCCCCAAGAACTCAGTAAACGCAGA ACTCCCTGGGCTGGACAC
-----+-----+-----+-----+-----+-----+-----+
17290      17310      17330
CATGCTGCGGGGAGGCAATAACCCACTCAGGATCACTGTGCCAACCTCCTGGACTCTTAT
-----+-----+-----+-----+-----+-----+-----+
17350      17370      17390
CACGTTGCTCAGCCCCAAAGATGGCCCCACACAGGGACCACCCCCTGGGCGGCGTTCACCC
-----+-----+-----+-----+-----+-----+-----+
17410      17430      17450
CATGGAGTTGGGCCCCACAGCCACTGGACAG CCTGGCCCCCTGGGTGAGTGAGGCACCAG
-----+-----+-----+-----+-----+-----+-----+
17470      17490      17510
GGGGAGGTGGAGAGGGAAGGGAGAAGGGAAGGGCTCATGCCTCCTGCCT CCTTCCAGATG
-----+-----+-----+-----+-----+-----+-----+
17530      17550      17570
GGCAGCACCCAGTCACCTTGAGTCCCCCTATGCCCCCTCCCCAGCCCCAGGGTCTCCTGCTG
-----+-----+-----+-----+-----+-----+-----+
17590      17610      17630
ACCATATTACACAACATTTACCCTCCACCATTTCTCCCAGACCCTGAGAACTCTCATGAGC
-----+-----+-----+-----+-----+-----+-----+
17650      17670      17690

```

FIG. 29

125/157

CCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCCCCAAGGTAGGCAGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
17710                      17730                      17750  
  
GACCTGGATTCAAAGCCTCCCCCTCTCATCGCCCACCCTCCCACCTCTCCCACCCCTCAG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
17770                      17790                      17810  
  
TTTGCTGCCCCCTAATCAGGTTTCTGGGCTCAGGTTATTATGGAAATGAGTTTATGACCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
17830                      17850                      17870  
  
CTTGGTTATCATGGAGACCAGGATGCTGGAAGCCCCTGGGCTGGGGAGGGGAGAAGCTGTG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
17890                      17910                      17930  
  
GCTTTTCCTGGATCACTGGTCCCTCACTGAGTGAGGATGGGCTCTCTGCCACACAGCTTGC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
17950                      17970                      17990  
  
AGCCTGGGGCCCCAGTCCTTAGGGGACAACATATCCTCCTCATCTCAGCAGATCAAGTC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
18010                      18030                      18050  
  
CAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAATGAACAGATTTAAAGAC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
18070                      18090                      18110  
  
AGGTGGCCACTGACAGCCACTCCAGGAACCTGAACTGCAGGGGCAGAGCCAGTGAATCAC  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
18130                      18150                      18170  
  
CGGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTTCTTCCCCGAGTGGAGCTTCGACCCA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
  
18190                      18210                      18230  
  
CCCACTCCAGGAACCCAGAGCCACATTAGAAGTTCTTGAGGGCTGGAGAACACTGCTGGG  
-----+-----+-----+-----+-----+-----+-----+-----+-----+

FIG. 29

126/157

|   |       |       |
|---|-------|-------|
| 18250   | 18270 | 18290 |
| CACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGGTCACACAGCCCCCTGACC         |       |       |
| -+-----+  |       |       |
| 18310   | 18330 | 18350 |
| TCCCTCACCACTGGAGGCTGGGTAGTGCTGGCCATCCCAAAAGGGCTCTGTCCTGGGAGT          |       |       |
| -+-----+  |       |       |
| 18370   | 18390 | 18410 |
| CTGGTGTGTCTCCTACATGCAATTTCCACGGACCCAGCTCTGTGGAGGGCATGACTGCTG          |       |       |
| -+-----+  |       |       |
| 18430   | 18450 | 18470 |
| GCCAGAAGCTAGTGGTCCTGGGGCCCTATGGTTCGACTGAGTCCACACTCCCCTGCAGCC          |       |       |
| -+-----+  |       |       |
| 18490   | 18510 | 18530 |
| TGGCTGGCCTCTGCAAA <u>CA</u> AACATAATTTGGGGACCTTCCTTCCTGTTTCTTCCCACCC  |       |       |
| -+-----+  |       |       |
| 18550   | 18570 | 18590 |
| TGTCTTCTCCCCTAGGTGGTTCCTGAGCCCCCACCCCCAATCCCAGTGCTACACCTGAGG          |       |       |
| -+-----+  |       |       |
| 18610   | 18630 | 18650 |
| TTCTGGAGCTCAGAACTCGACAGCCTCTCCCCCATTCTGTGTGTGTGCGGGGGACAGAGG          |       |       |
| -+-----+  |       |       |
| 18670   | 18690 | 18710 |
| GAACCATTTAAGAAAAAGATACCAAAGTAGAAGTCAAAGAAAGACATGTTGGCTATAGGC          |       |       |
| -+-----+  |       |       |
| 18730   | 18750 | 18770 |
| GTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAAGC <u>CG</u> GGGTAGGAGGATCACCAGA |       |       |
| -+-----+  |       |       |
| 18790   | 18810 | 18830 |

FIG. 29

GGCCAGCAGGTCCACACCAGCCTGGGCAACACAGCAAGACACCGCATCTACAGAAAAATT  
-----+-----+-----+-----+-----+-----+

18850 18870 18890

TTAAAATTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGCTGCTCAGGAGGCTGAAG  
-----+-----+-----+-----+-----+-----+

18910 18930 18950

CAGGAGGATCACTTGAGCCTGAGTTCAACACTGCAGTGAGCTATGGTGGCACCACTGCAC  
-----+-----+-----+-----+-----+-----+-----+

18970 18990 19010

TCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCTAAAATAAAATTTTAAAAAGACATATTA

-----+-----+-----+-----+-----+-----+-----+-----+

19030                      19050                      19070

ACTTGGACCTTGGTTAGTCCTTTCTGTATGTAAATTCAACCCATGGGGTGCCCTGAGGAC  
-----+-----+-----+-----+-----+

19090                      19110                      19130  
CCACACGGGGTGGTGTTGGCGGGGTGGTGGTTGGTGGGGTGGTGGCTGACGGGGTGGTG  
-----+-----+-----+-----+-----+-----+

19150                      19170                      19190

GCTGGCAGGCCGAGCCTAGATGGCAGCCAGAGCCCCAGGCATGTGTCTGGGCACAGGACG  
-----+-----+-----+-----+-----+

19210                      19230                      19250

GTGTTGCCTAGTTTGAACACCCTCTTTGCTCTGTCACTCCCTGGGCGTTCAC  
-----+-----+-----+-----+-----+

ATTCTCCCATTCGCTTCATGCAAGAGCTGCTGAGTGGCCTATATCAGCCAGCTGTTGCCGC  
-----+-----+-----+-----+-----+

19330 19350 19370  
ATAACAAAACCATCCCAAACTGAGTGCAGGGAGGCAACTTCACCTCGGGCTCCACTCCA

FIG. 29

128/157

```
-----+-----+-----+-----+-----+-----+
      19390              19410              19430
CAAGCCCAAGGGGCCAGGTGAGAGTGCTCTCTAAAGCCCCCTCCTGCCTCAGTTGTAGTT
-----+-----+-----+-----+-----+-----+
      19450              19470              19490
GCAAAATTTTAAATTTATGAAGGTGACTGATGACACAGAGGCCAATGCTGTTGAAATAAGT
-----+-----+-----+-----+-----+-----+
      19510              19530              19550
TATTACTCACAGTTTCCCACCATGCAGGGCCACAGTGGGGAGGCACTAGGTTTGGTCCAG
-----+-----+-----+-----+-----+-----+
      19570              19590              19610
GGACAGAATCAGGAGCGAGTGGAAGGCACAGGCCACAGCCCACAGTGCCGTTTCCACTGG
-----+-----+-----+-----+-----+-----+
      19630              19650              19670
GGAGGCAAGGCAGGCCAGGGGAAGAGGGTAGGATTGGCATTTTGAATCATTCTGGTGGGG
-----+-----+-----+-----+-----+-----+
      19690              19710              19730
TTTGGGGCGTGGGGTTGGGCTCTAATTGTCTGGGTAGGTGCCTGGCCCTGAGCTGGTTTA
-----+-----+-----+-----+-----+-----+
      19750              19770              19790
GGGCAGGGGAAATACTGGTTTTCGTATGTGAGAGTTCCTTGAAGGGGGTGGTTGGTGTATG
-----+-----+-----+-----+-----+-----+
      19810              19830              19850
GACTCAAGACTGGTCGGTTTGCATATGAAAGGCATGAGTTGTTTCTGATCTCCAGGAATC
-----+-----+-----+-----+-----+-----+
      19870              19890              19910
AAGCAGTTTCTCTCCAGCCAACAAGCCCCCAGATGTAAACCATCATAAATAG
-----+-----+-----+-----+-----+-----+
```

FIG. 29

19930 19950 19970  
AGAATCTAAGGCCAGGCATGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAG  
-----+-----+-----+-----+-----+-----+-----+  
19990 20010 20030  
GCGGGAGGATCATTGTGAGGTGAGAAAGTTGAGACCAGCCTGGCCAATGTGGTGAAACCCC  
-----+-----+-----+-----+-----+-----+-----+  
20050 20070 20090  
ATCTCTACTAAAAATACAAAATTAGCCCCGGTGTGGTGGCACGTGCCTATAATCCCAGCT  
-----+-----+-----+-----+-----+-----+-----+  
20110 20130 20150  
ACTCGGGAGGCTGCGGCAGGAGAATTGTTTGAACATGGGAGGTGGAGGTGTCAGTGAGCT  
-----+-----+-----+-----+-----+-----+-----+  
20170 20190 20210  
GAGATCGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCAAGACTCCGTCTCAAAAAAAAA  
-----+-----+-----+-----+-----+-----+-----+  
20230 20250 20270  
AAAAAAAAAAAAAGAGAGACTCTAAAAATACACGTTAATATACCTCCCCCGCTCTTACCCCT  
-----+-----+-----+-----+-----+-----+-----+  
20290 20310 20330  
TCAGGAGGGGGTGTCTAGACCCCCGCGGGAAGTCCAGCTACAGGGACCCCTGGGGAGGCCAA  
-----+-----+-----+-----+-----+-----+-----+  
20350 20370 20390  
CTCTGCCCTCTTGGCTAATCCCCAAGACTGCCCAGCACCCCTCCACCCCTTCTCCATTCC  
-----+-----+-----+-----+-----+-----+-----+  
20410 20430 20450  
AGTGGCGAACCCCTGGGGAGGCCACGTGGGAAGGAAAGAGGGCTCTAAGAGGGGAGGCCCC  
-----+-----+-----+-----+-----+-----+-----+  
20470 20490 20510  
AGACTGGGGGAGAGGCCTGTCTGGAGCCCAGGATCACCTGGCTGTGCTGCAGAACTGGAA  
-----+-----+-----+-----+-----+-----+-----+

FIG. 29

130/157

```
-----+-----+-----+-----+-----+-----+
      20530              20550              20570
AAGAGAAGCTCAGCAGAAAGGAGCTGGCATGGGGCCAACAGCAGAAAAGCAGGAGGCACG
-----+-----+-----+-----+-----+-----+

      20590              20610              20630
CAGAACTGACTGGGAAGCAGGAGGGTAGGCATGGACCCTGAGGCTGAGCAGGAGGTACTG
-----+-----+-----+-----+-----+-----+

      20650              20670              20690
AGGGGCAGAGTGGACGCTGAGCTGGGGGTAGCGAGCGAGCCCAGCTCAGCTGTGACGCCC
-----+-----+-----+-----+-----+-----+

      20710              20730              20750
TCTGTTTGGCCACCCAACTACCAGCTACTTGGGCTGCCCCGGGAGGAACTGGGCTTCCTC
-----+-----+-----+-----+-----+-----+

      20770              20790              20810
TGACATTCTGTGGCCTGCGGCCATCTGTGCACACCTTCTTCTCTCTGCCCCCTCCCTTGA
-----+-----+-----+-----+-----+-----+

      20830              20850              20870
CTTGTTGGCACCCACAGACAGGTGGGAGAGTGTACCTGCCCTGTGTGGTCAGAGCTTGGTT
-----+-----+-----+-----+-----+-----+

      20890              20910              20930
TTGAGTTTCCTTCCCTCACCCCTCTTTCTCCACACGCCAAAACACAAGAGGATGTGTC
-----+-----+-----+-----+-----+-----+

      20950              20970              20990
AGAGGCCTGTGAACCAGAGCAACTCCATCCTGAATAGGGGCTGAGCAAAATAAGGCTGAG
-----+-----+-----+-----+-----+-----+

      21010              21030              21050
ACCTACTGGGCTGCGTTTCCAGACAGTTACAGCATTCTGCGTCACAGGATGAGATAGGAG
-----+-----+-----+-----+-----+-----+
```

FIG. 29





132/157

```
-----+-----+-----+-----+-----+-----+
      21670              21690              21710
TTCTTTTATTCTTTTACTTTCTTAATAAATGTGCTTTCACTTTATGGACTCGTCTCAAAT
-----+-----+-----+-----+-----+-----+

      21730              21750              21770
TCTTTCTTGACGAGATCCAAGAACCCTCTCCTGGGGTCTGAATCTGGACCCCTTTCCGG
-----+-----+-----+-----+-----+-----+

      21790              21810              21830
TAACAGATGTCGTAGAGTGAAGCACAACCACTGCAGGGGCATCTTGGTTTACATTTTGCT
-----+-----+-----+-----+-----+-----+

      21850              21870              21890
TCAGCGGCCATGGTTAGCACAGCGGAAAGCACATCACAGTCTTCTGATTCATTAAAAAAA
-----+-----+-----+-----+-----+-----+

      21910              21930              21950
TTAGGAAATGGACCACCACAAACCACAGACAGATGTACTGAGACAGGATAGGTAGTCAAG
-----+-----+-----+-----+-----+-----+

      21970              21990              22010
AAAGTGACCATGTTCTAGGCGCGCAGCAGCAACTGTGGTGACCGTACAGTCAACAAGCCT
-----+-----+-----+-----+-----+-----+

      22030              22050              22070
CAGCACTGGCATTGCAATTGAGCTCATTCAAGCAAAGCTATCTTCAGCAGGGACTTCTCC
-----+-----+-----+-----+-----+-----+

      22090              22110              22130
CTCTAGGCAGCAAGCGCATTTTTATTTTACCTGTCCTCAAACCTGATCCTTTGCTCCTTAT
-----+-----+-----+-----+-----+-----+

      22150              22170              22190
AACAGTAAGGAACACACCCCTGTGTGGAGATTTAAGATGCTAATGAGGCCAAGCGCAGTT
-----+-----+-----+-----+-----+-----+
```

FIG. 29

22210                      22230                      22250  
GCTCACGTCTGTAATTCAGCACTTTGGGAGGCAGAGGTGGCGGGCTCACTTGAGGTTAG  
-----+-----+-----+-----+-----+-----+  
  
22270                      22290                      22310  
AAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTTGCTCTACTAAAAATACAAAAAT  
-----+-----+-----+-----+-----+-----+  
  
22330                      22350                      22370  
TAGCCGGGCATGGTGGCGGGCGTCTGTAATCCCAGCTACCTGGGAGGCTGAGGCAGAAGA  
-----+-----+-----+-----+-----+-----+  
  
22390                      22410                      22430  
ATCGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGCCAAGATCGTGCCACTGCACTCCAG  
-----+-----+-----+-----+-----+-----+  
  
22450                      22470                      22490  
CCTGAGGGAGAGAGAGAGCAAGACATCGTTTTTGT TTGTTGTTGTTGTTGTTGTTGTT  
-----+-----+-----+-----+-----+-----+  
  
22510                      22530                      22550  
TTTTAAAAAAAAGTCAAGACAAATCATAGTGGGGGCTTTTCTGGTCACTTTTTTAAATCTT  
-----+-----+-----+-----+-----+-----+  
  
22570                      22590                      22610  
AGTGTTGAGACTTTATT TGAGACAGGGCCTTACTCTGTTGCCAGGTTGGATGAGATTTT  
-----+-----+-----+-----+-----+-----+  
  
22630                      22650                      22670  
TAACCTCAATATTTACTTATAGAATAACTTTTGGTTAGTCAAACAATGCTGTGTCTCA  
-----+-----+-----+-----+-----+-----+  
  
22690                      22710                      22730  
TTCTGATCAGAATAAAACATCAGACAACTCAAGAGAAACATTCTGCAAAATAACTGGCCA  
-----+-----+-----+-----+-----+-----+  
  
22750                      22770                      22790  
GGATTCTTCAAAAGTGTC AAGGGTAAAGATAAGGAAAAGATGAAGGA ACTCCCAGATTGA

FIG. 29

134/157

-----+-----+-----+-----+-----+-----+  
22810 22830 22850  
GAGAATAAGGAGACAACTGTGATGTGGGATCCTAGAATGGATCTTGGAACAGAAAAGGA  
-----+-----+-----+-----+-----+-----+  
22870 22890 22910  
CATTAGTGGA AAAATGAGAAATGCAAAACAGTCTACAGTTTCGTTAACAGGATTGTACCA  
-----+-----+-----+-----+-----+-----+  
22930 22950 22970  
AGGTTAGTTTCCTAGCTGTAATGATTGGACTATGATTAAGTAAGATGGACCATCAGGGGA  
-----+-----+-----+-----+-----+-----+  
22990 23010 23030  
AGCTGGGTGAAGGGTGTAAGGAAAATGCTTACATTTTCCAACTTTCTGCAAGTCTAAAA  
-----+-----+-----+-----+-----+-----+  
23050 23070 23090  
TTAGTCAACAATAAGAAGTTTAAAATAGGCCAGGCATGGTGGCTCACACCTGTCATCCTA  
-----+-----+-----+-----+-----+-----+  
23110 23130 23150  
GCACTTTGAGAGGCCGAGGTGGGAGGATGGTTTGAGCCCAGGAGTTCAAGACCAGCCTGC  
-----+-----+-----+-----+-----+-----+  
23170 23190 23210  
GCAATAGAGCGAGACCCCAACTCTATTCAAAAAAATTTTTTAAGTTTAAAATAGAATTA  
-----+-----+-----+-----+-----+-----+  
23230 23250 23270  
TATAAAAAGGAAAAGAAAATGCTGTTTCATAGCGTTCCTAGTTTAGCATGGGAGAGAC  
-----+-----+-----+-----+-----+-----+  
23290 23310 23330  
CAGGTCTCCCTGGGTGGTTGTCTGTGTGTTGCTGGGTGTGCGTGCCAGGGCTAGTGTGTT  
-----+-----+-----+-----+-----+-----+

FIG. 29

135/157

23350 23370 23390

GGGGTCCGTCTAGGCACATT CAGGCGCCGAATCCCGTGGCTCCCAGGTTTACCTGACGGT  
-----+-----+-----+-----+-----+-----+

23410 23430 23450

GCAGCCTGGGGTGGAGACTTAATGAGGGCGGGGAGTTGCTGCAGCAAAGGCTCCTCCCAG  
-----+-----+-----+-----+-----+-----+

23470 23490 23510

GGGTATCAGCGCAGACAGCTGGGTTTTCACTGTGCTCCTGCTCCAGAGGCACTAGGAAGG  
-----+-----+-----+-----+-----+-----+

23530 23550 23570

GGGCGCCTATCAGACTAGGACTCTGCCAGCCATCCTTCTCTGTTGAAGGTCCAGC  
-----+-----+-----+-----+-----+-----+

FIG. 29

136/157

```

      10              30              50
CAGCTATGGGCTGGAGGCCCCGGAGAGCTCGGGGGACCCCGTTGCTGCTGCTGCTACTAC
-----+-----+-----+-----+-----+-----+-----+
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuLeuLeuLeuLeuLeuL

      70              90              110
TGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGACATATCCCTGGGCAGC
-----+-----+-----+-----+-----+-----+
euLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyHisIleProGlyGlnP

      130             150             170
CAGTCACCCCGCACTGGGTCCTGGATGGACAACCCTGGCGCACCGTCAGCCTGGAGGAGC
-----+-----+-----+-----+-----+-----+
roValThrProHisTrpValLeuAspGlyGlnProTrpArgThrValSerLeuGluGluP

      190             210             230
CGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAGGCCAGGAGCTCCTGC
-----+-----+-----+-----+-----+-----+
roValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluGlyGlnGluLeuLeuL

      250             270             290
TTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATAGAAACCCACTACGGCC
-----+-----+-----+-----+-----+-----+
euGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIleGluThrHisTyrGlyP

      310             330             350
CAGATGGGCAGCCAGTGCTGGCCCCCAACCACACGGATCATTGCCACTACCAAGGGC
-----+-----+-----+-----+-----+-----+
roAspGlyGlnProValValLeuAlaProAsnHisThrAspHisCysHisTyrGlnGlyA

      370             390             410
GAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCC
-----+-----+-----+-----+-----+-----+
rgValArgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyL

      430             450             470
TGATCACCCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCA
-----+-----+-----+-----+-----+-----+
euIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerL

```

FIG. 30

137/157

```

      490              510              530
AGGACTTCTCAACCCACGAGATCTTTTCGGATGGAGCAGCTGCTCACCTGGAAAGGAACCT
-----+-----+-----+-----+-----+-----+
ysAspPheSerThrHisGluIlePheArgMetGluGlnLeuLeuThrTrpLysGlyThrC

      550              570              590
GTGGCCACAGGGATCCTGGGAACAAAGCGGGCATGACCAGCCTTCCTGGTGGTCCCCAGA
-----+-----+-----+-----+-----+-----+
ysGlyHisArgAspProGlyAsnLysAlaGlyMetThrSerLeuProGlyGlyProGlns

      610              630              650
GCAGGGGCAGGCGAGAAGCGCGCAGGACCCGGAAGTACCTGGAAGTGTACATTGTGGCAG
-----+-----+-----+-----+-----+-----+
erArgGlyArgArgGluAlaArgArgThrArgLysTyrLeuGluLeuTyrIleValAlaA

      670              690              710
ACCACACCCTGTTCTTGAAGTCTGGCACCAGAACTTGAACCACACCAAACAGCGTCTCCTGG
-----+-----+-----+-----+-----+-----+
spHisThrLeuPheLeuThrArgHisArgAsnLeuAsnHisThrLysGlnArgLeuLeuG

      730              750              770
AAGTCGCCAACTACGTGGACCAGCTTCTCAGGACTCTGGACATTCAGGTGGCGCTGACCG
-----+-----+-----+-----+-----+-----+
luValAlaAsnTyrValAspGlnLeuLeuArgThrLeuAspIleGlnValAlaLeuThrG

      790              810              830
GCCTGGAGGTGTGGACCGAGCGGGACCGCAGCCGCGTCACGCAGGACGCCAACGCCACGC
-----+-----+-----+-----+-----+-----+
lyLeuGluValTrpThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrL

      850              870              890
TCTGGGCCTTCCTGCAGTGGCGCCGGGGGCTGTGGGCGCAGCGGCCCCACGACTCCGCGC
-----+-----+-----+-----+-----+-----+
euTrpAlaPheLeuGlnTrpArgArgGlyLeuTrpAlaGlnArgProHisAspSerAlaG

      910              930              950
AGCTGCTCACGGGCCGCGCCTTCCAGGGCGCCACAGTGGGCCTGGCGCCCGTCGAGGGCA
-----+-----+-----+-----+-----+-----+
lnLeuLeuThrGlyArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyM

```

FIG. 30

138/157

|   |      |      |
|---|------|------|
| 970   | 990  | 1010 |
| TGTGCCGCGCCGAGAGCTCGGGAGGCGTGAGCACGGACCACTCGGAGCTCCCCATCGGCG<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>etCysArgAlaGluSerSerGlyGlyValSerThrAspHisSerGluLeuProIleGlyA  |      |      |
| 1030  | 1050 | 1070 |
| CCGCAGCCACCATGGCCCATGAGATCGGCCACAGCCTCGGCCTCAGCCACGACCCCGACG<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>laAlaAlaThrMetAlaHisGluIleGlyHisSerLeuGlyLeuSerHisAspProAspG  |      |      |
| 1090  | 1110 | 1130 |
| GCTGCTGCGTGGAGGCTGCGGCCGAGTCCGGAGGCTGCGTCATGGCTGCGGCCACCGGGC<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>lyCysCysValGluAlaAlaAlaGluSerGlyGlyCysValMetAlaAlaAlaThrGlyH  |      |      |
| 1150  | 1170 | 1190 |
| ACCCGTTTCCGCGCGTGTTTCAGCGCCTGCAGCCGCCGCCAGCTGCGCGCCTTCTTCCGCA<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>isProPheProArgValPheSerAlaCysSerArgArgGlnLeuArgAlaPhePheArgL |      |      |
| 1210  | 1230 | 1250 |
| AGGGGGGCGGCGCTTGCTCTCCAATGCCCGGACCCCGGACTCCCGGTGCCGCCGGCGC<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>ysGlyGlyGlyAlaCysLeuSerAsnAlaProAspProGlyLeuProValProProAlaL    |      |      |
| 1270  | 1290 | 1310 |
| TCTGCGGGAACGGCTTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGEGCCAGGAGT<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>euCysGlyAsnGlyPheValGluAlaGlyGluGluCysAspCysGlyProGlyGlnGluC |      |      |
| 1330  | 1350 | 1370 |
| GCCGCGACCTCTGCTGCTTTGCTCACAACCTGCTCGCTGCGCCCGGGGGCCAGTGCGCCC<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>ysArgAspLeuCysCysPheAlaHisAsnCysSerLeuArgProGlyAlaGlnCysAlaH  |      |      |
| 1390  | 1410 | 1430 |
| ACGGGGACTGCTGCGTGCGCTGCTGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCA<br>-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+<br>isGlyAspCysCysValArgCysLeuLeuLysProAlaGlyAlaLeuCysArgGlnAlaM   |      |      |
| 1450  | 1470 | 1490 |

FIG. 30

139/157

```

TGGGTGACTGTGACCTCCCTGAGTTTTGCACGGGCACCTCCTCCCACTGTCCCCAGACG
-----+-----+-----+-----+-----+-----+-----+
etGlyAspCysAspLeuProGluPheCysThrGlyThrSerSerHisCysProProAspV

      1510              1530              1550
TTTACCTACTGGACGGCTCACCTGTGCCAGGGGCAGTGGCTACTGCTGGGATGGCGCAT
-----+-----+-----+-----+-----+-----+-----+
alTyrLeuLeuAspGlySerProCysAlaArgGlySerGlyTyrCysTrpAspGlyAlaC

      1570              1590              1610
GTCCACCGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCCTGGCTCCCACCCAGCTCCCG
-----+-----+-----+-----+-----+-----+-----+
ysProThrLeuGluGlnGlnCysGlnGlnLeuTrpGlyProGlySerHisProAlaProG

      1630              1650              1670
AGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGATGCTCATGGAACTGCGGCCAGGACA
-----+-----+-----+-----+-----+-----+-----+
luAlaCysPheGlnValValAsnSerAlaGlyAspAlaHisGlyAsnCysGlyGlnAspS

      1690              1710              1730
GCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCC
-----+-----+-----+-----+-----+-----+-----+
erGluGlyHisPheLeuProCysAlaGlyArgAspAlaLeuCysGlyLysLeuGlnCysG

      1750              1770              1790
AGGGTGGAAGCCCAGCCTGCTCGCACCCGCACATGGTGCCAGTGGACTCTACCGTTCACC
-----+-----+-----+-----+-----+-----+-----+
lnGlyGlyLysProSerLeuLeuAlaProHisMetValProValAspSerThrValHisL

      1810              1830              1850
TAGATGGCCAGGAAGTGACTTGTCGGGGAGCCTTGGCACTCCCCAGTGCCCAGCTGGACC
-----+-----+-----+-----+-----+-----+-----+
euAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLeuAspL

      1870              1890              1910
TGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTGCCAGA
-----+-----+-----+-----+-----+-----+-----+
euLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCysGlnS

      1930              1950              1970

```

FIG. 30



140/157

```

GCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTGCCACA
-----+-----+-----+-----+-----+-----+
erArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCysHisS

      1990              2010              2030
GCCACGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCCT
-----+-----+-----+-----+-----+
erHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaProProp

      2050              2070              2090
TCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGAAAACC
-----+-----+-----+-----+-----+
heCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGluAsnH

      2110              2130              2150
ATGACACCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGGGGCCG
-----+-----+-----+-----+-----+
isAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGlyAlaG

      2170              2190              2210
GCCTGGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCT
-----+-----+-----+-----+-----+
lyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyC

      2230              2250              2270
GCAGAAGGGACCCTGCGTGACAGTGGCCCCAAAGATGGCCCACACAGGGACCACCCCTGG
-----+-----+-----+-----+-----+
ysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuG

      2290              2310              2330
GCGGCGTTCACCCCATGGAGTTGGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCTGGACC
-----+-----+-----+-----+-----+
lyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspP

      2350              2370              2390
CTGAGAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTG
-----+-----+-----+-----+-----+
roGluAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProA

      2410              2430              2450
ACCCCCAAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAA

```

FIG. 30

141/157

```
-----+-----+-----+-----+-----+-----+
spProGlnAspGlnValGlnMetProArgSerCysLeuTrpEnd

      2470              2490              2510
TGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAAGTTGAACTGCAGGGG
-----+-----+-----+-----+-----+-----+

      2530              2550              2570
CAGAGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGGAAAGTTTCTTCCCCG
-----+-----+-----+-----+-----+-----+

      2590              2610              2630
AGTGGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACATAGAAAGTTCCTGAGGGC
-----+-----+-----+-----+-----+-----+

      2650              2670              2690
TGGAGAACACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGG
-----+-----+-----+-----+-----+-----+

      2710              2730              2750
TCACACAGCCCCCTGACCTCCCTCACCAGTGGAGGCTGGGTAGTGCTGGCCATCCCAAAG
-----+-----+-----+-----+-----+-----+

      2770              2790              2810
GGCTCTGTCTGGGAGTCTGGTGTGTCTCTACATGCAATTTCCACGGACCCAGCTCTGT
-----+-----+-----+-----+-----+-----+

      2830              2850              2870
GGAGGGCATGACTGCTGGCCAGAAGCTAGTGGTCCTGGGGCCCTATGGTTCGACTGAGTC
-----+-----+-----+-----+-----+-----+

      2890              2910              2930
CACACTCCCCTGCAGCCTGGCTGGCCTCTGCAAACAAACATAATTTGGGGACCTTCCTT
-----+-----+-----+-----+-----+-----+

      2950              2970              2990
CCTGTTTCTTCCCAACCCTGTCTTCTCCCCTAGGTGGTTCTTGAGCCCCCACCCCCAATCC
-----+-----+-----+-----+-----+-----+
```

FIG. 30

142/157

```

      3010              3030              3050
      .               .               .
CAGTGCTACACCTGAGGTTCTGGAGCTCAGAATCTGACAGCCTCTCCCCATTCTGTGTG
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3070              3090              3110
      .               .               .
TGTCGGGGGGACAGAGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAAAGAAAG
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3130              3150              3170
      .               .               .
ACATGTTGGCTATAGGCGTGGTGGCTCATGCCTATAATCCCAGCACCTTTGGGAAGC_GGG
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3190              3210              3230
      .               .               .
GTAGGAGGATCACCAGAGGCCAG_CAGGTCCACACCAGCCTGGGCAACACAGCAAGACACC
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3250              3270              3290
      .               .               .
GCATCTACAG_AAAAAATTTTAAAAATTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGC
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3310              3330              3350
      .               .               .
TGCTCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTGAGTTCAACACTGCAGTGAGCTA
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3370              3390              3410
      .               .               .
TGGTGGCACCACTGCACCTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCTAAAATAAAT
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3430              3450              3470
      .               .               .
TTTAAAAAGACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+
      .               .               .
      3490
      .               .               .
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+

```

FIG. 30

143/157

10 30 50  
CGGGCACGGGTCGGCCGCAATCCAGCCTGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG

70 90 110  
AGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCGTTGC  
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

130 150 170  
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGAC  
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

190 210 230  
ATATCCCTGGGCAGCCAGTCACCCGCACTGGGTCCTGGATGGACAACCCTGGCGCACCG  
isIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThrV

250 270 290  
TCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAAG  
alSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGluG

310 330  
GCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGC  
lyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArg

FIG. 31

144/157

10 30 50  
CGGGCACGGGTCGGCCGCAATCCAGCCTGGGCGGAGCCGGAGTTGCGAGCCGCTGCCTAG

70 90 110  
AGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCCGGAGAGCTCGGGGGACCCCGTTGC  
MetGlyTrpArgProArgArgAlaArgGlyThrProLeuL

130 150 170  
TGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGAC  
euLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGlyH

190 210  
ATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGAC  
isIleProGlyGlnProValThrProHisTrpValLeuAspGly

FIG. 32

145/157

10 30 50  
GCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCC  
MetGlyTrpArgProArgArgAlaArgGlyThrP

70 90 110  
CGTTGCTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTC  
roLeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuG

130 150 170  
AAGGACATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCTCGGATGGACAACCCTGGC  
lnGlyHisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpA

190 210 230  
GCACCGTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGG  
rgThrValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluA

250 270 290  
CTGAAGGCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGAT  
laGluGlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyT

310 330 350  
ACATAGAAACCCACTACGGCCAGATGGGCAGCCAGTGGTGTGGCCCCCAACCACACGG  
yrIleGluThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrV

370 390 410  
TGAGATGCTTCCATGGGCTCTGGGATGCACCGCCAGAGGATCATTGCCACTACCAAGGGC  
alArgCysPheHisGlyLeuTrpAspAlaProProGluAspHisCysHisTyrGlnGlyA

430 450 470  
GAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCC  
rgValArgGlyPheProAspSerTrpValValLeuCysThrCysSerGlyMetSerGlyL

490 510 530  
TGATCACCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCA  
euIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrpProProArgGlySerL

550  
AGGACTTCTCAACCCACGAGAT  
ysAspPheSerThrHisGlu

FIG. 33

146/157

10 30 50  
GAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTG  
MetGlyTrpArgProArgArgAlaArgGlyThrProLeu

70 90 110  
CTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGA  
LeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGly

130 150 170  
CATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCTGGCGCACC  
HisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThr

190 210 230  
GTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAA  
ValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGlu

250 270 290  
GGCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCATGGCCTGATCACCCCTCAGCAGGAAT  
GlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisGlyLeuIleThrLeuSerArgAsn

310 330 350  
GCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACTTCTCAACCCACGAG  
AlaSerTyrTyrLeuArgProTrpProProArgGlySerLysAspPheSerThrHisGlu

AT

FIG. 34

147/157

10 30 50  
GAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGTTG  
MetGlyTrpArgProArgArgAlaArgGlyThrProLeu

70 90 110  
CTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGA  
LeuLeuLeuLeuLeuLeuLeuLeuLeuTrpProValProGlyAlaGlyValLeuGlnGly

130 150 170  
CATATCCCTGGGCAGCCAGTCACCCCGCACTGGGTCCTGGATGGACAACCCTGGCGCACC  
HisIleProGlyGlnProValThrProHisTrpValLeuAspGlyGlnProTrpArgThr

190 210 230  
GTCAGCCTGGAGGAGCCGGTCTCGAAGCCAGACATGGGGCTGGTGGCCCTGGAGGCTGAA  
ValSerLeuGluGluProValSerLysProAspMetGlyLeuValAlaLeuGluAlaGlu

250 270 290  
GGCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGCTGCTGGCCCCAGGATACATA  
GlyGlnGluLeuLeuLeuGluLeuGluLysAsnHisArgLeuLeuAlaProGlyTyrIle

310 330 350  
GAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGCTGGCCCCCAACCACACGGATCAT  
GluThrHisTyrGlyProAspGlyGlnProValValLeuAlaProAsnHisThrAspHis

370 390 410  
TGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCCTCTGCACCTGC  
CysHisTyrGlnGlyArgValArgGlyPheProAspSerTrpValValLeuCysThrCys

430 450 470  
TCTGGGATGAGTGGCCTGATCACCCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGG  
SerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSerTyrTyrLeuArgProTrp

490 510  
CCACCCCGGGGCTCCAAGGACTTCTCAACCCACGAGAT  
ProProArgGlySerLysAspPheSerThrHisGlu

FIG. 35



148/157

10 30 50  
CTGGCCCCAGGATACATAGAAACCCACTACGGCCCAGATGGGCAGCCAGTGGTGCTGGCC  
LeuAlaProGlyTyrIleGluThrHisTyrGlyProAspGlyGlnProValValLeuAla

70 90 110  
CCCAACCACACGGATCATTGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGG  
ProAsnHisThrAspHisCysHisTyrGlnGlyArgValArgGlyPheProAspSerTrp

130 150 170  
GTAGTCCTCTGCACCTGCTCTGGGATGAGTGGCCTGATCACCTCAGCAGGAATGCCAGC  
ValValLeuCysThrCysSerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSer

190 210 230  
TATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACTTCTCAACCCACGAGATCTTT  
TyrTyrLeuArgProTrpProProArgGlySerLysAspPheSerThrHisGluIlePhe

250 270 290  
CGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGGCCACAGGGATCCTGGGAACAAA  
ArgMetGluGlnLeuLeuThrTrpLysGlyThrCysGlyHisArgAspProGlyAsnLys

310 330 350  
G  
GCGGGCATGACCAGCCTTCCTGGTGGTCCCCAGAGCAGGGGCAGGCGAAAAGCGCGCAGG  
AlaGlyMetThrSerLeuProGlyGlyProGlnSerArgGlyArgArgLysAlaArgArg  
Glu

370 390 410  
ACCCGGAAGTACCTGGAAGTGTACATTGTGGCAGACCACACCCTGTTCTTGACTCGGCAC  
ThrArgLysTyrLeuGluLeuTyrIleValAlaAspHisThrLeuPheLeuThrArgHis

430 450 470  
CGAAACTTGAACCACACCAAACAGCGTCTCCTGGAAGTCGCCAACTACGTGGACCAGCTT  
ArgAsnLeuAsnHisThrLysGlnArgLeuLeuGluValAlaAsnTyrValAspGlnLeu

490  
CTCAGGACTCTGGACATTCAGGTGGC  
LeuArgThrLeuAspIleGlnVal

FIG. 36

149/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGGCCTGGCTCCCACCCAGCTCCCGAGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGA  
pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs

130 150 170  
TGCTCATGGAACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA  
pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs

190 210 230  
TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAGCCCAGCCTGCTCGCACCGCACAT  
pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe

250 270 290  
GGTGCCAGTGGACTCTACCGTTCACCTAGATGGCCAGGAAGTGACTTGTCGGGGAGCCTT  
tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe

310 330 350  
GGCACTCCCCAGTGCCCAGCTGGACCTGCTTGGCCTGGGCTGGTAGAGCCAGGCACCCCA  
uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl

370 390 410  
GTGTGGACCTAGAATGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGC  
nCysGlyProArgMetValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAl

430 450 470  
TCCACCCTTCTGTGACAAGCCAGGCTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGC  
aProProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAl

490  
TGAAAACCATGACACCTTCCTGC  
aGluAsnHisAspThrPheLeu

FIG. 37

150/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGSCCTGGCTCCCACCCAGCTCCCGAGGCTGTTTCCAGGTGGTGAACCTCTGCGGGAGA  
pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs

130 150 170  
TGCTCATGGAAACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA  
pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs

190 210 230  
TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCCAGCCTGCTCGCACCGCACAT  
pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe

250 270 290  
GGTGCCAGTGGACTCTACCGTTCACCTAGATGGCCAGGAAGTGACTTGTCGGGGAGCCTT  
tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe

310 330 350  
GGCACTCCCCAGTGGCCAGCTGGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCA  
uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl

370 390 410  
GTGTGGACCTAGAATGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCT  
nCysGlyProArgMetValCysGlnSerArgArgCysArgLysAsnAlaPheGlnGluLe

430 450 470  
TCAGCGCTGCCTGACTGCCTGCCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCA  
uGlnArgCysLeuThrAlaCysHisSerHisGlyValCysAsnSerAsnHisAsnCysHi

490 510 530  
CTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGA  
sCysAlaProGlyTrpAlaProProPheCysAspLysProGlyPheGlyGlySerMetAs

550 570  
CAGTGGCCCTGTGCAGGCTGAAAACCATGACACCTTCCTGC  
pSerGlyProValGlnAlaGluAsnHisAspThrPheLeu

FIG. 38

151/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGGCCTGATGGCCAGGAAGTGACTTGTCGGGGAGCCTTGCGCACTCCCCAGTGCCCAGCT  
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

130 150 170  
GGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTG  
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

190 210 230  
CCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG  
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

250 270 290  
CCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC  
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

310 330 350  
ACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA  
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

370  
AAACCATGACACCTTCCTGC  
uAsnHisAspThrPheLeu

FIG. 39

152/157

10 30 50  
GGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGGGCTGCAG  
AlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTrpGlyCysAr

70 90 110  
AAGGGACCCCTGCGTGCGTGGCCCCAAAGATGGCCCACACAGGGACCACCCCTGGGCGG  
gArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisProLeuGlyGl

130 150 170  
CGTTCACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCTGGACCCCTGA  
yValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLeuAspProGl

190 210 230  
GAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCC  
uAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSerProAspPr

250 270 290  
CCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCCTAAAAT  
oGlnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

310  
GAACAGATTTAAAGACAGGTGGCC

FIG. 40

153/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGGCCTGATGGCCAGGAAGTGACTTGTCTGGGGAGCCTTGGCACTCCCCAGTGCCCAGCT  
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

130 150 170  
GGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTG  
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

190 210 230  
CCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG  
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

250 270 290  
CCACAGCCACGGGGTTTGAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC  
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

310 330 350  
ACCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA  
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

370 390 410  
AAACCATGACACCTTCTGCTGGCCATGCTCCTCAGCGTCTGCTGCCTCTGCTCCCAGG  
uAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGl

430 450 470  
GGCCGGCCTGGCCTGGTGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTG  
yAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTr

490 510 530  
GGGCTGCAGAAGGGACCCTGCGTGCAGTGGCCCCAAAGATGGCCACACAGGGACCACCC  
pGlyCysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisPr

550 570 590  
CCTGGGCGGCGTTACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCT  
oLeuGlyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLe

610 630 650  
GGACCTGAGAACTCTCATGAGCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTC  
uAspProGluAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSe

670 690 710  
GCCTGACCCCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGC  
rProAspProGlnAlaAspGlnValGlnMetProArgSerCysLeuTrpEnd

730 750  
TCCTAAATGAACAGATTAAAGACAGGTGGCC

FIG. 41

154/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGGCCTGGCTCCCACCCAGCTCCCGAGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGA  
pGlyProGlySerHisProAlaProGluAlaCysPheGlnValValAsnSerAlaGlyAs

130 150 170  
TGCTCATGGAAACTGCGGCCAGGACAGCGAGGGCCACTTCCTGCCCTGTGCAGGGAGGGA  
pAlaHisGlyAsnCysGlyGlnAspSerGluGlyHisPheLeuProCysAlaGlyArgAs

190 210 230  
TGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCCAGCCTGCTCGCACCGCACAT  
pAlaLeuCysGlyLysLeuGlnCysGlnGlyGlyLysProSerLeuLeuAlaProHisMe

250 270 290  
GGTGCCAGTGGACTCTACCGTTCACCTAGATGGCCAGGAAGTGACTTGTGCGGGAGCCTT  
tValProValAspSerThrValHisLeuAspGlyGlnGluValThrCysArgGlyAlaLe

310 330 350  
GGCACTCCCCAGTGCCAGCTGGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCA  
uAlaLeuProSerAlaGlnLeuAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGl

370 390 410  
GTGTGGACCTAGAATGGTGTGCCAGAGCAGGCGCTGCAGGAAGAATGCCTTCCAGGAGCT  
nCysGlyProArgMetValCysGlnSerArgArgCysArgLysAsnAlaPheGlnGluLe

430 450 470  
TCAGCGCTGCCTGACTGCCTGCCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCA  
uGlnArgCysLeuThrAlaCysHisSerHisGlyValCysAsnSerAsnHisAsnCysHi

490 510 530  
CTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGA  
sCysAlaProGlyTrpAlaProProPheCysAspLysProGlyPheGlyGlySerMetAs

550 570 590  
CAGTGGCCCTGTGCAGGCTGAAAACCATGACACCTTCCTGCTGGCCATGCTCCTCAGCGT  
pSerGlyProValGlnAlaGluAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerVa

610 630 650  
CCTGCTGCCTCTGCTCCCAGGGGCCGGCCTGGCCTGGTGTGCTACCGACTCCAGGAGC  
lLeuLeuProLeuLeuProGlyAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAl

670 690 710

FIG. 42

155/157

CCATCTGCAGCGATGCAGCTGGGGCTGCAGAAAGGGACCCTGCGTGCCAGTGGCCCCAAAGA  
aHisLeuGlnArgCysSerTrpGlyCysArgArgAspProAlaCysSerGlyProLysAs

730

750

770

TGGCCACACAGGGACCACCCCTGGGCGGCCTTCACCCCATGGAGTTGGGCCCCACAGC  
pGlyProHisArgAspHisProLeuGlyGlyValHisProMetGluLeuGlyProThrAl

790

810

830

CACTGGACAGCCCTGGCCCTGGACCCTGAGAACTCTCATGAGCCCAGCAGCCACCCTGA  
aThrGlyGlnProTrpProLeuAspProGluAsnSerHisGluProSerSerHisProGl

850

870

890

GAAGCCTCTGCCAGCAGTCTCGCCTGACCCCAAGATCAAGTCCAGATGCCAAGATCCTG  
uLysProLeuProAlaValSerProAspProGlnAspGlnValGlnMetProArgSerCy

910

930

950

CCTCTGGTGAGAGGTAGCTCCTAAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGC  
sLeuTrpEnd

970

990

1010

CACTCCAGGAACCTTGAAGTGCAGGGGCAGAGCCAGTGAATCACCGGACCTCCAGCACCTG

1030

1050

1070

CAGGCAGCTTGGAAGTTTCTTCCCCGAGTGGAGCTTCGACCCACCCACTCCAGGAACCCA

1090

1110

1130

GAGCCACATTAGAAGTTCCTGAGGGCTGGAGAACACTGCTGGGCACACTCTCCAGCTCAA

1150

TAAACCATCAGTCC

rt772

FIG. 42



156/157

10 30 50  
CAGTGGCTACTGCTGGGATGGCGCATGTCCCACGCTGGAGCAGCAGTGCCAGCAGCTCTG  
SerGlyTyrCysTrpAspGlyAlaCysProThrLeuGluGlnGlnCysGlnGlnLeuTr

70 90 110  
GGGGCCTGATGGCCAGGAAGTGACTTGTCTGGGGAGCCTTGGCACTCCCCAGTGCCAGCT  
pGlyProAspGlyGlnGluValThrCysArgGlyAlaLeuAlaLeuProSerAlaGlnLe

130 150 170  
GGACCTGCTTGGCCTGGGCCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAATGGTGTG  
uAspLeuLeuGlyLeuGlyLeuValGluProGlyThrGlnCysGlyProArgMetValCy

190 210 230  
CCAGAGCAGGCCTGCAGGAAGAATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTG  
sGlnSerArgArgCysArgLysAsnAlaPheGlnGluLeuGlnArgCysLeuThrAlaCy

250 270 290  
CCACAGCCACGGGGTTTGCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCC  
sHisSerHisGlyValCysAsnSerAsnHisAsnCysHisCysAlaProGlyTrpAlaPr

310 330 350  
ACCCTTCTGTGACAAGCCAGGCTTTGGTGGCAGCATGGACAGTGGCCCTGTGCAGGCTGA  
oProPheCysAspLysProGlyPheGlyGlySerMetAspSerGlyProValGlnAlaGl

370 390 410  
AAACCATGACACCTTCCTGCTGGCCATGCTCCTCAGCGTCCTGCTGCCTCTGCTCCCAGG  
uAsnHisAspThrPheLeuLeuAlaMetLeuLeuSerValLeuLeuProLeuLeuProGl

430 450 470  
GGCCGGCCTGGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTG  
yAlaGlyLeuAlaTrpCysCysTyrArgLeuProGlyAlaHisLeuGlnArgCysSerTr

490 510 530  
GGGCTGCAGAAAGGGACCCTGCGTGCACTGGCCCCAAAGATGGCCCACACAGGGACCACCC  
pGlyCysArgArgAspProAlaCysSerGlyProLysAspGlyProHisArgAspHisPr

550 570 590

FIG. 43

157/157

CCTGGGCGGCGTTTACCCCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCCCT  
oLeuGlyGlyValHisProMetGluLeuGlyProThrAlaThrGlyGlnProTrpProLe

610

630

650

GGACCCTGAGAACTCTCATGAGCCCAGCAGCCACCCTGAGAAGCCTCTGCCAGCAGTCTC  
uAspProGluAsnSerHisGluProSerSerHisProGluLysProLeuProAlaValSe

670

690

710

GCCTGACCCCCAAGATCAAGTCCAGATGCCAAGATCCTGCCTCTGGTGAGAGGTAGCTCC  
rProAspProGlnAspGlnValGlnMetProArgSerCysLeuTrpEnd

730

750

770

TAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAAGTTGAACTGC

790

810

830

AGGGGCAGAGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGGAAGTTTCTT

850

870

890

CCCCGAGTGGAGCTTCGACCCACCCACTCCAGGAACCCAGAGCCACATTAGAAGTTCCTG

910

930

950

AGGGCTGGAGAACTGCTGGGCACACTCTCCAGCTCAATAAACCATCAGTCC

FIG. 43

## SEQUENCE LISTING

&lt;110&gt; GENOME THERAPEUTICS CORPORATION

&lt;120&gt; NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES AND OBESITY

&lt;130&gt; 2976-4039PCT

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 363

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 3626

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

|            |            |            |            |             |             |      |
|------------|------------|------------|------------|-------------|-------------|------|
| cgggcacggg | tgggccgcaa | tccagcctgg | gcggagccgg | agttgcgagc  | cgctgcctag  | 60   |
| aggccgagga | gctcacagct | atgggctgga | ggccccggag | agctcggggg  | accccgttgc  | 120  |
| tgtctgtgt  | actactgtct | ctgtctctgg | cagtgccagg | cgccgggggtg | cttcaaggac  | 180  |
| atatccctgg | gcagccagtc | accccgcact | gggtcctgga | tggacaaccc  | tggtgcaccg  | 240  |
| tcagcctgga | ggagccggtc | tcgaagccag | acatggggct | ggtggccctg  | gaggctgaag  | 300  |
| gccaggagct | cctgcttgag | ctggagaaga | accacaggct | gctggcccca  | ggatacatag  | 360  |
| aaacccacta | cggcccagat | gggcagccag | tggtgctggc | ccccaaccc   | acggtgagat  | 420  |
| gcttccatgg | gctctgggat | gcaccgccag | aggatcattg | ccactaccaa  | gggcgagtaa  | 480  |
| ggggcttccc | cgactcctgg | gtagtctct  | gcacctgtct | tgggatgagt  | ggcctgatca  | 540  |
| ccctcagcag | gaatgccagc | tattatctgc | gtccctggcc | accccggggc  | tccaaggact  | 600  |
| tctcaaccca | cgagatcttt | cggatggagc | agctgtctac | ctggaaagga  | acctgtggcc  | 660  |
| acagggatcc | tgggaacaaa | gcgggcatga | ccagccttcc | tggtgggtccc | cagagcaggg  | 720  |
| gcagggcaga | agcgcgcagg | acccggaagt | acctggaact | gtacattgtg  | gcagaccaca  | 780  |
| ccctgtttct | gactcggcac | cgaaacttga | accacaccaa | acagcgtctc  | ctggaagtcc  | 840  |
| ccaactacgt | ggaccagctt | ctcaggactc | tggacattca | ggtggcgctg  | accggcctgg  | 900  |
| aggtgtggac | cgagcggggc | cgcagcccg  | tcacgcagga | cgccaaacgc  | acgctctggg  | 960  |
| ccttcctgca | gtggcgcccg | gggctgtggg | cgcagcggcc | ccacgactcc  | gcgcagctgc  | 1020 |
| tcacggggcc | cgccttccag | ggcgccacag | tgggcctggc | gcccgtcgag  | ggcatgtgcc  | 1080 |
| gcgccgagag | ctcgggaggc | gtgagcacgg | accactcgga | gtccccatc   | ggcgccgcag  | 1140 |
| ccaccatggc | ccatgagatc | ggccacagcc | tcggcctcag | ccacgacccc  | gacggctgct  | 1200 |
| gcgtggaggc | tgcggccgag | tccggaggct | gcgtcatggc | tgcggccacc  | gggcacccgt  | 1260 |
| ttccgcgcgt | gttcagcgcc | tgcagccgcc | gccagctgcg | cgccttcttc  | cgcaaggggg  | 1320 |
| gcggcgcttg | cctctccaat | gccccggacc | ccggactccc | ggtgcgcgcg  | gcgctctgcg  | 1380 |
| ggaacggctt | cgtggaagcg | ggcgaggagt | gtgactgcgg | ccctggccag  | gagtgcgcgcg | 1440 |
| acctctgctg | ctttgtctac | aactgctcgc | tgcgcccggg | ggcccagtcg  | gcccacgggg  | 1500 |
| actgctgcgt | gcgctgcctg | ctgaagccgg | ctggagcgct | gtgcgcgcag  | gccatgggtg  | 1560 |
| actgtgacct | ccctgagttt | tgcacgggca | cctcctccca | ctgtccccc   | gacgtttacc  | 1620 |

2/154

|             |             |            |             |             |             |      |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| tactggacgg  | ctcaccctgt  | gccaggggca | gtggctactg  | ctgggatggc  | gcattgtcca  | 1680 |
| cgctggagca  | gcagtgccag  | cagctctggg | ggcctggctc  | ccaccagct   | cccagggcct  | 1740 |
| gtttccaggt  | ggtgaactct  | gcgggagatg | ctcatggaaa  | ctgcggccag  | gacagcgagg  | 1800 |
| gccacttct   | gcccgtgtga  | gggagggatg | ccctgtgtgg  | gaagctgcag  | tgccaggggtg | 1860 |
| gaaagcccag  | cctgctcgca  | ccgcacatgg | tgccagtggg  | ctctaccgtt  | cacctagatg  | 1920 |
| gccaggaagt  | gacttgtcgg  | ggagccttgg | cactccccag  | tgcccagctg  | gacctgcttg  | 1980 |
| gcctgggect  | ggtagagcca  | ggcaccagct | gtggacctag  | aatggtgtgc  | cagagcaggc  | 2040 |
| gctgcaggaa  | gaatgccttc  | caggagcttc | agcgtgcct   | gactgcctgc  | cacagccacg  | 2100 |
| gggtttgcaa  | tagcaaccat  | aactgccact | gtgctccagg  | ctgggctcca  | cccttctgtg  | 2160 |
| acaagccagg  | ctttggtggc  | agcatggaca | gtggccctgt  | gcaggctgaa  | aaccatgaca  | 2220 |
| ccttctgtct  | ggccatgctc  | ctcagcgtcc | tgtgtcctct  | gctcccaggg  | ggcggcctgg  | 2280 |
| ccctggtgtg  | ctaccgactc  | ccaggagccc | atctgcagcg  | atgcagctgg  | ggctgcagaa  | 2340 |
| gggacctgtc  | gtgcagtggc  | cccaaagatg | gccacacag   | ggaccacccc  | ctgggctggcg | 2400 |
| ttcaccccat  | ggagtgtggc  | cccacagcca | ctggacagcc  | ctggcccttg  | gaccttgaga  | 2460 |
| actctcatga  | gcccagcagc  | caccttgaga | agcctctgcc  | agcagtctcg  | cctgaccccc  | 2520 |
| aagcagatca  | agtccagatg  | ccaagatcct | gcctctgggtg | agaggttagct | cctaaaaatga | 2580 |
| acagatttaa  | agacaggtgg  | ccactgacag | ccactccagg  | aacttgaact  | gcagggggcag | 2640 |
| agccagtga   | tcaccggacc  | tccagcacct | gcaggcagct  | tggaaagttc  | ttccccgagt  | 2700 |
| ggagcttcga  | cccacccact  | ccaggaaccc | agagccacat  | tagaagttcc  | tgagggtctgg | 2760 |
| agaacactgc  | tgggcacact  | ctccagctca | ataaaccatc  | agtcccagaa  | gcaaagggtca | 2820 |
| cacagccct   | gacctccctc  | accagtggag | gctgggtagt  | gctggccatc  | ccaaaagggc  | 2880 |
| tctgtcctgg  | gagtctgggtg | tgtctcctac | atgcaatttc  | cacggaccca  | gctctgtgga  | 2940 |
| gggcatgact  | gctggccaga  | agctagtggg | cctggggccc  | tatgggttcga | ctgagtccac  | 3000 |
| actccctgc   | agcctggctg  | gcctctgcaa | acaaacataa  | ttttggggac  | cttcttctct  | 3060 |
| gtttcttccc  | accctgtctt  | ctcccctagg | tggttctctga | gccccccacc  | ccaatcccag  | 3120 |
| tgctacacct  | gagggttctgg | agctcagaat | ctgacagcct  | ctccccatt   | ctgtgtgtgt  | 3180 |
| cgggggggaca | gaggggaacca | tttaagaaaa | gataccaaag  | tagaagtcaa  | aagaaaagaca | 3240 |
| tgttggtat   | aggcgtgggtg | gctcatgcct | ataatcccag  | cactttggga  | agccggggta  | 3300 |
| ggaggtacac  | cagaggccag  | caggtccaca | ccagcctggg  | caacacagca  | agacaccgca  | 3360 |
| tctacagaaa  | aatttttaaaa | ttagctgggc | gtggtggtgt  | gtacctgtag  | gcctagctgc  | 3420 |
| tcaggaggct  | gaagcaggag  | gatcacttga | gcctgagttc  | aacactgcag  | tgagctatgg  | 3480 |
| tggcaccact  | gcactccagc  | ctgggtgaca | gagcaagacc  | ctgtctctaa  | aataaaatttt | 3540 |
| aaaaagacat  | aaaaaaaaaa  | aaaaaaaaaa | aaaaaaaaaa  | aaaaaaaaaa  | aaaaaaaaaa  | 3600 |
| aaaaaaaaaa  | aaaaaaaaaa  | aaaaaa     |             |             |             | 3626 |

<210> 2  
 <211> 227  
 <212> DNA  
 <213> Homo sapiens

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| <400> 2    |            |            |            |            |            |     |
| accgggcacg | ggtcggccgc | aatccagcct | gggcggagcc | ggagtgtgca | gcccgtgcct | 60  |
| agaggccgag | gagctcacag | ctatgggctg | gaggccccgg | agagctcggg | ggaccccgtt | 120 |
| gctgctgctg | ctactactgc | tgtgtctctg | gccagtgcc  | ggcgccgggg | tgcttcaagg | 180 |
| acatatccct | gggcagccag | tcaccccga  | ctgggtcctg | gatggac    |            | 227 |

<210> 3  
 <211> 3509  
 <212> DNA  
 <213> Homo sapiens

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| <400> 3    |            |            |            |            |            |     |
| cagctatggg | ctggaggccc | cggagagctc | gggggacccc | gttgctgctg | ctgtacttac | 60  |
| tgctgtgctg | ctggccagtg | ccaggcgccg | gggtgcttca | aggacatatc | cctgggcagc | 120 |
| cagtcacccc | gcactgggtc | ctggatggac | aaccctggcg | caccgtcagc | ctggaggagc | 180 |
| cggctctgaa | gccagacatg | gggctggtgg | ccctggaggc | tgaaggccag | gagctcctgc | 240 |

|             |             |             |            |             |             |      |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| ttgagctgga  | gaagaaccac  | aggctgctgg  | ccccaggata | catagaaacc  | cactacggcc  | 300  |
| cagatgggca  | gccagtgggt  | ctggccccc   | accacacgga | tcattgccac  | taccaagggc  | 360  |
| gagtaagggg  | cttccccgac  | tcctgggtag  | tcctctgcac | ctgctctggg  | atgagtggcc  | 420  |
| tgatcacccct | cagcagggaat | gccagctatt  | atctgcgtcc | ctggccaccc  | cggggctcca  | 480  |
| aggacttctc  | aaccacagag  | atctttcggg  | tggagcagct | gctcacctgg  | aaaggaaacct | 540  |
| gtggccacac  | ggatcctggg  | aacaaagcgg  | gcatgaccag | ccttcctggg  | gggtcccaga  | 600  |
| gcagggggcag | gcgagaagcg  | cgcaggaccc  | ggaagtacct | ggaactgtac  | attgtggcag  | 660  |
| accacaccct  | gttcttgact  | cggcaccgaa  | acttgaacca | caccaaacag  | cgtctcctgg  | 720  |
| aagtcgcca   | ctacgtggac  | cagcttctca  | ggactctgga | cattcagggtg | gcgctgaccg  | 780  |
| gcctggagggt | gtggaccgag  | cgggaccgca  | gccgcgtcac | gcaggacgcc  | aacgccacgc  | 840  |
| tctgggcctt  | cctgcagtgg  | cggcgggggc  | tgtgggcgca | gcggccccc   | gactccgcgc  | 900  |
| agctgctcac  | gggccgcgcc  | ttccaggggc  | ccacagtggg | cctggcgccc  | gtcgagggca  | 960  |
| tgtgccgcgc  | cgagagctcg  | ggaggcgtga  | gcacggacca | ctcggagctc  | cccatcggcg  | 1020 |
| ccgcagccac  | catggcccat  | gagatcggcc  | acagcctcgg | cctcagccac  | gaccccgacg  | 1080 |
| gctgctgcgt  | ggaggctgcg  | gccgagtcgg  | gaggctgcgt | catggctgcg  | gccaccgggc  | 1140 |
| acccgtttcc  | gcgcgtgttc  | agcgccctgca | gccgcgcgca | gctgcgcgcc  | ttcttccgca  | 1200 |
| agggggggcgg | cgtttgcctc  | tccaatgccc  | cggaccccg  | actcccgggtg | ccgcggcgcc  | 1260 |
| tctgcgggaa  | cggcttcgtg  | gaagcggggc  | aggagtgtga | ctgcggccct  | ggccaggagt  | 1320 |
| gcgcgcacct  | ctgctgcttt  | gctcacaact  | gctcgtcgcg | cccggggggc  | cagtgcgcct  | 1380 |
| acggggactg  | ctgcgtgcgc  | tgcttgcgtga | agccggctgg | agcgtgtgtc  | cggcaggcca  | 1440 |
| tgggtgactg  | tgacctccct  | gagttttgca  | cgggcacctc | ctcccactgt  | ccccagacg   | 1500 |
| tttacctact  | ggacggctca  | ccctgtgcca  | ggggcagttg | ctactgctgg  | gatggcgcat  | 1560 |
| gtcccacgct  | ggagcagcag  | tgccagcagc  | tctggggggc | tggtcccac   | ccagctcccg  | 1620 |
| aggcctgttt  | ccagggtggtg | aactctgcgg  | gagatgctca | tggaaactgc  | ggccaggaca  | 1680 |
| gcgaggggcca | cttccctgccc | tgtgcaggga  | gggatgccct | gtgtgggaag  | ctgcagtgcc  | 1740 |
| aggggtggaaa | gcccagcctg  | ctcgcaccgc  | acatggtgcc | agtggactct  | accgttcacc  | 1800 |
| tagatggcca  | ggaagtgact  | tgtcggggag  | ccttggcact | ccccagtgcc  | cagctggacc  | 1860 |
| tgcttggcct  | gggcctggta  | gagccaggca  | cccagttggg | acctagaatg  | gtgtgccaga  | 1920 |
| gcaggcgctg  | caggaagaat  | gccttccagg  | agcttcagcg | ctgcctgact  | gctgccaca   | 1980 |
| gccacgggggt | ttgcaatagc  | aaccataact  | gccactgtgc | tccaggctgg  | gctccacct   | 2040 |
| tctgtgacaa  | gccaggcttt  | ggtggcagca  | tggacagtgg | ccctgtgcag  | gctgaaaacc  | 2100 |
| atgacacctt  | cctgctggcc  | atgctcctca  | gcgtcctgct | gcctctgctc  | ccagggggccg | 2160 |
| gcctggcctg  | gtgttgctac  | cgactcccag  | gagcccatct | gcagcgatgc  | agctggggct  | 2220 |
| gcagaaggga  | ccctgcgtgc  | agtggcccca  | aagatggccc | acacaggggac | cacccctgg   | 2280 |
| gcggcggttca | ccccatggag  | ttggggcccca | cagccactgg | acagccctgg  | ccctggacc   | 2340 |
| ctgagaactc  | tcatgagccc  | agcagccacc  | ctgagaagcc | tctgccagca  | gtctgcctg   | 2400 |
| acccccaaga  | tcaagtccag  | atgccaaagt  | cctgcctctg | gtgagaggta  | gctcctaaaa  | 2460 |
| tgaacagatt  | taaagacagg  | tggccactga  | cagccactcc | aggaaactga  | actgcagggg  | 2520 |
| cagagccagt  | gaatcacccg  | acctccagca  | cctgcaggca | gcttgggaagt | ttcttccccg  | 2580 |
| agtggagctt  | cgacccaccc  | actccaggaa  | cccagagcca | cattagaagt  | tctgaggggc  | 2640 |
| tggagaacac  | tgtctgggcac | actctccagc  | tcaataaacc | atcagtccca  | gaagcaaagg  | 2700 |
| tcacacagcc  | cctgacctcc  | ctcaccagtg  | gaggctgggt | agtgtctggc  | atcccaaaag  | 2760 |
| ggctctgtcc  | tgggagtctg  | gtgtgtctcc  | tacatgcaat | ttccacggac  | ccagctctgt  | 2820 |
| ggagggcatg  | actgctggcc  | agaagctagt  | ggtcctgggg | ccctatggtt  | cgactgagtc  | 2880 |
| cacactcccc  | tgcagcctgg  | ctggcctctg  | caaacaaaca | taattttggg  | gaccttcctt  | 2940 |
| cctgtttctt  | cccacctgt   | cttctccct   | aggtggttcc | tgagccccc   | cccccaatcc  | 3000 |
| cagtgtctaca | cctgagggtt  | tggagctcag  | aatctgacag | cctctcccc   | attctgtgtg  | 3060 |
| tgtcgggggg  | acagagggaa  | ccattttaaga | aaagatacca | aagtagaagt  | caaaagaaag  | 3120 |
| acatgtttggc | tataggcgtg  | gtggctcatg  | cctataatcc | cagcactttg  | ggaagccggg  | 3180 |
| gtaggaggat  | caccagaggc  | cagcaggtcc  | acaccagcct | gggcaacaca  | gcaagacacc  | 3240 |
| gcactctacag | aaaaatttta  | aaattagctg  | ggcgtgggtg | tgtgtacctg  | taggcctagc  | 3300 |
| tgctcaggag  | gctgaagcag  | gaggatcact  | tgagcctgag | ttcaacactg  | cagttagcta  | 3360 |
| tggtggcacc  | actgcactcc  | agcctgggtg  | acagagcaag | accctgtctc  | taaaataaat  | 3420 |
| tttaaaaaaga | cataaaaaaa  | aaaaaaaaaa  | aaaaaaaaaa | aaaaaaaaaa  | aaaaaaaaaa  | 3480 |
| aaaaaaaaaa  | aaaaaaaaaa  | aaaaaaaaaa  |            |             |             | 3509 |

4/154

&lt;211&gt; 826

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

```

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu
 1           5           10           15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln
      20           25           30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly
      35           40           45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp
      50           55           60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu
 65           70           75           80

Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His
      85           90           95

Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Val
      100          105          110

Arg Cys Phe His Gly Leu Trp Asp Ala Pro Pro Glu Asp His Cys His
      115          120          125

Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys
      130          135          140

Thr Cys Ser Gly Met Ser Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser
      145          150          155          160

Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr
      165          170          175

His Glu Ile Phe Arg Met Glu Gln Leu Leu Thr Trp Lys Gly Thr Cys
      180          185          190

Gly His Arg Asp Pro Gly Asn Lys Ala Gly Met Thr Ser Leu Pro Gly
      195          200          205

Gly Pro Gln Ser Arg Gly Arg Arg Glu Ala Arg Arg Thr Arg Lys Tyr
      210          215          220

Leu Glu Leu Tyr Ile Val Ala Asp His Thr Leu Phe Leu Thr Arg His
      225          230          235          240

Arg Asn Leu Asn His Thr Lys Gln Arg Leu Leu Glu Val Ala Asn Tyr
      245          250          255

Val Asp Gln Leu Leu Arg Thr Leu Asp Ile Gln Val Ala Leu Thr Gly
      260          265          270

Leu Glu Val Trp Thr Glu Arg Asp Arg Ser Arg Val Thr Gln Asp Ala
      275          280          285

```

5/154

Asn Ala Thr Leu Trp Ala Phe Leu Gln Trp Arg Arg Gly Leu Trp Ala  
 290 295 300  
 Gln Arg Pro His Asp Ser Ala Gln Leu Leu Thr Gly Arg Ala Phe Gln  
 305 310 315 320  
 Gly Ala Thr Val Gly Leu Ala Pro Val Glu Gly Met Cys Arg Ala Glu  
 325 330 335  
 Ser Ser Gly Gly Val Ser Thr Asp His Ser Glu Leu Pro Ile Gly Ala  
 340 345 350  
 Ala Ala Thr Met Ala His Glu Ile Gly His Ser Leu Gly Leu Ser His  
 355 360 365  
 Asp Pro Asp Gly Cys Cys Val Glu Ala Ala Ala Glu Ser Gly Gly Cys  
 370 375 380  
 Val Met Ala Ala Ala Thr Gly His Pro Phe Pro Arg Val Phe Ser Ala  
 385 390 395 400  
 Cys Ser Arg Arg Gln Leu Arg Ala Phe Phe Arg Lys Gly Gly Gly Ala  
 405 410 415  
 Cys Leu Ser Asn Ala Pro Asp Pro Gly Leu Pro Val Pro Pro Ala Leu  
 420 425 430  
 Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Pro  
 435 440 445  
 Gly Gln Glu Cys Arg Asp Leu Cys Cys Phe Ala His Asn Cys Ser Leu  
 450 455 460  
 Arg Pro Gly Ala Gln Cys Ala His Gly Asp Cys Cys Val Arg Cys Leu  
 465 470 475 480  
 Leu Lys Pro Ala Gly Ala Leu Cys Arg Gln Ala Met Gly Asp Cys Asp  
 485 490 495  
 Leu Pro Glu Phe Cys Thr Gly Thr Ser Ser His Cys Pro Pro Asp Val  
 500 505 510  
 Tyr Leu Leu Asp Gly Ser Pro Cys Ala Arg Gly Ser Gly Tyr Cys Trp  
 515 520 525  
 Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys Gln Gln Leu Trp Gly  
 530 535 540  
 Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe Gln Val Val Asn Ser  
 545 550 555 560  
 Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp Ser Glu Gly His Phe  
 565 570 575  
 Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly Lys Leu Gln Cys Gln  
 580 585 590

6/154

Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met Val Pro Val Asp Ser  
           595                                600                                605  
 Thr Val His Leu Asp Gly Gln Glu Val Thr Cys Arg Gly Ala Leu Ala  
           610                                615                                620  
 Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val Glu Pro  
  625                                630                                635                                640  
 Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg Cys Arg  
                                 645                                650                                655  
 Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys His Ser  
                                 660                                665                                670  
 His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro Gly Trp  
                                 675                                680                                685  
 Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met Asp Ser  
           690                                695                                700  
 Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala Met Leu  
  705                                710                                715                                720  
 Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala Trp Cys  
                                 725                                730                                735  
 Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp Gly Cys  
                                 740                                745                                750  
 Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His Arg Asp  
           755                                760                                765  
 His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr Ala Thr  
           770                                775                                780  
 Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro Ser Ser  
  785                                790                                795                                800  
 His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln Ala Asp  
                                 805                                810                                815  
 Gln Val Gln Met Pro Arg Ser Cys Leu Trp  
                                 820                                825

<210> 5  
 <211> 207433  
 <212> DNA  
 <213> Homo sapiens

<400> 5  
 gctctaataa atttgccggcc gctaatacga ctactatag ggagaggatc cgcggaattc 60  
 ccccatgtgc catgtccacg agatggttga acagatgaga caacatgggt gtgcagcttc 120  
 tctctttttt tttttttcag acagagtctc actctgtagc ccaggctgga gtgcagtggc 180  
 gcaatttcag ctactgcaa cctccgcctc ccaggctctg attcaagcag ttctcctgcc 240  
 tcagcctcct gactagctgg gattacagga acgcgccact atgccagct aatttttcta 300  
 ggttttagtag agacgggggtt tcaccatgtt gaccaggctg gtcttgaatt cctgaccttg 360



7/154

|            |             |             |             |             |             |      |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| tgatccgccc | gcctcggcct  | cccgaagtgc  | tgagattaca  | ggcatgagcc  | accgcatccg  | 420  |
| gccatgcagc | ttctcttttc  | tagagttaac  | aggagatgcg  | gcagggtctgt | agtagcccgag | 480  |
| aaattctcag | acttgccactc | ttgaattttac | acgcagtgac  | aaaatgcagc  | tcagccctaag | 540  |
| gcctcctggt | cagttctgtg  | gtgctgccta  | cttcctgtgc  | ccaaattagc  | cttgcggtta  | 600  |
| taactaggag | gaatgaacttg | gtgtcctaca  | atatttttggc | tcctgggctt  | ctttgggtgtt | 660  |
| cattctatta | ctagttgatt  | ttttttttct  | tggaaagaaa  | agtattcaaa  | gagccagaat  | 720  |
| ttactttcat | tatatcttac  | agttgttaaa  | ttaacctatc  | cttgttttca  | ctttctgtgt  | 780  |
| acttctttct | ttgggggtaca | aaacagtgtt  | tcttttagatt | gtctattttct | aaatattatt  | 840  |
| ttaaccagaa | cataccaaag  | atatttttct  | gccagtttca  | ttcttcattt  | tacttcttaa  | 900  |
| ccattgttac | gattttttttt | ttaacttcgg  | tctcagattt  | tgctcaattg  | aagttgtccc  | 960  |
| agttcatgag | attttgtttt  | ctttgcagct  | cttgaccatg  | acagatgtga  | ccagcacaca  | 1020 |
| gattgttaca | gctgcacagc  | caacaccaat  | gactgccact  | gggtgcaatga | ccattgtgtc  | 1080 |
| cccaggaacc | acagctgctc  | agaaggccag  | gtcagaaggc  | gtttctttaa  | gatttcagaa  | 1140 |
| aaatcccaat | ttgtcatagg  | tttagctttt  | atagtgtata  | tgggtataaat | aatggcccag  | 1200 |
| agttactttt | caaattgggtt | tctattttgga | ttttattatc  | cctgagggtt  | tccttttagga | 1260 |
| agagatgttc | tgtatatattc | aagggtccta  | ctagccctgg  | gaatttgtgt  | attgtgtttt  | 1320 |
| agaagaaggt | aggactgttg  | tccctgggca  | tggctgtgac  | taagcactgg  | atccttggtt  | 1380 |
| tgaggcatct | gatagtgacc  | tcactttacc  | agtaccaggt  | tttgataaag  | ttaggggttg  | 1440 |
| agagtgaagc | aggtgtacag  | cggcccatat  | tgaactgtgg  | aatgacaga   | ctttgcaaaa  | 1500 |
| tctccttggt | ttatatcttg  | gtttggcata  | atctcactgc  | tcatcatata  | tagattttta  | 1560 |
| aatttaagtt | ataagatgac  | ccaggctttt  | atagtttttg  | acagcttaca  | agactttttt  | 1620 |
| ttttgtcagt | ctcatcagt   | tcataaaata  | gaaaactttc  | agaacttttg  | agtgaagcatt | 1680 |
| tgaaaagcac | caagttcaac  | attcccattt  | tcagatgag   | cagggtgagg  | catggtgttt  | 1740 |
| aaaagagtgg | gctgcttctc  | tgccagttaa  | aagcatgtcc  | ttgagcagaa  | cattccattg  | 1800 |
| cagtgttccc | cactgcagag  | ctactgcaac  | ttatctatct  | atctgtctgt  | ctgtctgtct  | 1860 |
| atctatctat | ctatctatct  | atctatctat  | ctatctatct  | gtctatctat  | ttatttgcag  | 1920 |
| tttgcttctc | caattagatt  | ttgttgtcta  | taatatcctg  | ttggctctca  | ccataagttc  | 1980 |
| tctctgtgga | tacaaaagaa  | aagcttttgt  | ccactgtgtt  | ctagatctcc  | attttttaggt | 2040 |
| atgagaattg | cccccaaggat | aaccccatgt  | actactgtaa  | caagaagacc  | agctgcagga  | 2100 |
| gctgtgccct | ggaccagaac  | tgccagtggg  | agccccggaa  | tcaggagtgc  | attgccctgc  | 2160 |
| ccggtaggcc | ttgcagggtc  | atcttggtgt  | gtgtgggtcc  | attacttcag  | cctgcttccc  | 2220 |
| ccaacactgt | gcagcctaag  | ttgaacctag  | cagaggggaa  | gagctaattc  | tgtccattca  | 2280 |
| ccccccacac | gagtattatg  | ggcttttttg  | tttttaacta  | aaatacagtt  | cttaagtatt  | 2340 |
| tgttcctact | gtcctttgaa  | ataaagtga   | acatcctttg  | ctgctctgta  | gaattgagtg  | 2400 |
| acagcttggt | tctgtatcac  | tgagcctgct  | ccttgctctc  | ttcccatcca  | cctttgataa  | 2460 |
| cctgggacta | gaccatgatg  | tctcagcagt  | cagtctgctg  | agcactttat  | ggagagtact  | 2520 |
| tcttattaac | cactgggatt  | taattgttgg  | cacctgctaa  | tgggccttct  | ctgagaagga  | 2580 |
| gaggatagat | acttctgtca  | gcagcacctt  | ttaggggtga  | tctccagccc  | tgaaaacctc  | 2640 |
| aatatatcct | gcttctgagg  | ttcaggatga  | tgaactcagg  | gcctgagacc  | agccagccat  | 2700 |
| tgatatattt | tggaccaggg  | tgggtccagaa | aggggaactt  | ccttgctctgt | gcacaacatc  | 2760 |
| ttgacttggt | caggggaagt  | ggtttggacc  | agggcctttt  | tgaatgtccc  | ttggaagtgt  | 2820 |
| ttgattcagc | ttgcagagtg  | ggactccttt  | ctatatctca  | gtgtgtctca  | aattttaagg  | 2880 |
| tacaggaaaa | gacctgggaa  | tcttgcgaaa  | ggcatattct  | gggtgcaagt  | tggggtgggg  | 2940 |
| cctagagtgt | gcattttatc  | tagtcgactg  | ctacactgtg  | aggagcaagt  | ctttgtctgt  | 3000 |
| tcttaaatga | tctctttccc  | atggtacctt  | tcttttatct  | cagtgaactgt | tactgttaat  | 3060 |
| gaacattggt | gatgtctcca  | aagtactttg  | gttctggtga  | agttgctttg  | ttctttcatt  | 3120 |
| gtcttccagg | aagcattcat  | agcttctgtg  | cagtaccttg  | tgtgggttca  | ggatgatcac  | 3180 |
| aggtagcaga | ttacaagctt  | gtcttgtatg  | ctatagccat  | atcacttggg  | ttgtttctca  | 3240 |
| agaaggacct | tctcaccttg  | cttttgggat  | gctttgtaca  | cttgattgta  | ccttccacct  | 3300 |
| gatgatatga | aaacagtgca  | gcttttggag  | actatagatt  | tgttaattcc  | ttgattcatt  | 3360 |
| tccattcttg | cagtttttac  | cccagccctc  | caatatgcac  | attcatttgt  | ctgctcttca  | 3420 |
| cttaggattt | tagttttcta  | attgttcttc  | agaagggaagt | gtaccagtct  | aatattggca  | 3480 |
| ccaaactggt | gttttcatct  | aagacatagg  | ataagtgacc  | tcagaatatg  | cttttttagga | 3540 |
| tccgggagat | atcaccagta  | aacattttta  | aattcttgta  | ttctgcattt  | ggctcctaat  | 3600 |
| aatgtgtcag | aggctcccac  | atcctaata   | agtacctaga  | atttaaatta  | gaaaggccat  | 3660 |
| ttcgggtatc | agtaatttga  | actcataata  | cagttagttt  | gtctgatttc  | taaaattctt  | 3720 |
| ttctttctct | tttcccttta  | atgaaagaaa  | atatctgtgg  | cattggctgg  | catttggttg  | 3780 |
| gaaactcatg | tttgaaaatt  | actactgcca  | aggagaatta  | tgacaatgct  | aaattgttct  | 3840 |

|            |             |            |             |             |             |      |
|------------|-------------|------------|-------------|-------------|-------------|------|
| gtaggaacca | caatgccctt  | ttggcttctc | ttacaaccca  | gaagaaggta  | gaatttgtcc  | 3900 |
| ttaagcagct | gcgaataatg  | cagtcacctc | agagcatggg  | gagttaaaat  | cctcaaaact  | 3960 |
| taagtttctg | gttatccacc  | tttctaccaa | gggcatgact  | gcagcttgca  | tgtggaaagg  | 4020 |
| tgtggatatg | tgtaacgtgc  | ttggcaagaa | ggggagtgtc  | ggtgaacgca  | gcctgagggg  | 4080 |
| ctgtgggttt | gtgctgtcag  | agtctcttcc | tcttaaaatt  | tttaataact  | tgtatatata  | 4140 |
| agatctatga | ataattatat  | gggggatgaa | ttgtaacatg  | tatatgtgta  | cataatctgg  | 4200 |
| tgacatcagt | agattatttc  | atacctgttt | tacctctgga  | ttctgctagg  | ggagaaagag  | 4260 |
| aggtcactga | taattagcta  | ggttggatta | agccacctga  | gttccttgga  | gttaagggtat | 4320 |
| tataatagt  | cataagactg  | tataattacc | actaagaagt  | gtacatctca  | gctggatgtg  | 4380 |
| gtggctcaca | cctgtaatcc  | cagcactttg | ggaggctgag  | gtgggtggat  | cgctgaggt   | 4440 |
| caggagttca | agaccagcct  | gaccaatata | gtaaaacccc  | gtctctacta  | aaaatacaaa  | 4500 |
| aattagccag | gcatggtggt  | gtgcgcctgt | aatcccagct  | actcaggagg  | cggaggcagg  | 4560 |
| agaattgctt | gaaccagga   | ggtaggaggt | gcagtgaagt  | gagattgctc  | caccgcactc  | 4620 |
| cagcctgggt | gacagagcga  | tactccgtct | caaaacaaaa  | aaagaacagc  | aaaaaaaaaga | 4680 |
| aatgtacatc | tccttgggtc  | tttcatagcc | cagccatctc  | aaaagaagag  | agcaccttct  | 4740 |
| tgtcaagagt | ttctaagccc  | agaaaggctc | aagtctctct  | cttgctccacc | cagtgtctct  | 4800 |
| agggggctta | tagtcaatat  | tccatgatca | cattttgtca  | tttttagtct  | ggagtcataa  | 4860 |
| attgtgatcc | taccgtcagt  | taagtagact | catagaacaa  | agctctttca  | gcagtttcag  | 4920 |
| ctgtggtaca | gaaacgttta  | gtggaaatgt | tcttaccagg  | cagggaggag  | ttgagggcaa  | 4980 |
| cacttctctg | ggtagagcct  | ccttcatgtg | aaggtatgga  | aatgtggcct  | gggtctcgcc  | 5040 |
| tgccgttgge | ctctgtgtac  | cacctatagg | gcattcttga  | gactggtagg  | aggtgccctg  | 5100 |
| tatttagttt | tctccaatta  | gtcccttttt | tcagtgaac   | ttagatgggg  | tatggacact  | 5160 |
| caaacattgg | tgacatatcc  | ttagtgtgtt | tacctcaggc  | tactgtgacc  | acatttggtg  | 5220 |
| tttcataata | ttttgatagc  | tttttcagat | ttcagaatct  | ccattggtga  | ctgtctctgt  | 5280 |
| tgtttctctt | tccatgtcca  | aatgtgggtc | tcttccagca  | ttccatcctt  | ggctggcagc  | 5340 |
| tgacctttcc | catagttatt  | cactccctca | gaaaatggat  | ggcaccagc   | tttgcttatt  | 5400 |
| accctgggtg | ctctaacaat  | gactctcgag | tccacggaag  | ttaaaagggt  | tcaaccagggt | 5460 |
| ggccacagat | atacttctgg  | taccttttct | ctcttcttta  | ggttttctaa  | ctctaaacgt  | 5520 |
| ttctgggtat | tctaactctg  | tgtggccacg | tttatgaaac  | agaaattcac  | agtcttaatg  | 5580 |
| ataagactga | cagatgagag  | acaactgaag | tgtaatgtcc  | ttccacagct  | atacctctag  | 5640 |
| atgtagccca | gtttagtagg  | cccagatttt | tatggaaaaa  | caagaaagga  | cacctagcct  | 5700 |
| aacccttagg | acagaggctg  | gctggtagaa | agctggaagg  | aggtgacccc  | tgcagggtgag | 5760 |
| aaggagtgaa | ttaggatgtc  | gaagacggaa | gggttttctg  | tgatttaatt  | agtgtcccca  | 5820 |
| tctgtgagat | gtagagggag  | atgattaagg | gagtggctct  | ttgagtgagc  | tgcaggtaag  | 5880 |
| tttgcatggg | tggtgtcagg  | ctggatggga | ggggattttt  | atagttgagc  | ctcagggaagg | 5940 |
| aaccaaggcc | agaccctgcg  | aggccatggc | tacaatagta  | atggatttga  | attgtatcgt  | 6000 |
| gaaggcaaa  | aaattattga  | agggccttta | aaaagtattt  | ttaattgtct  | ttctttcttt  | 6060 |
| tggatctgtt | tatttttagg  | tcttttagta | ggcatgtgta  | ttttttctct  | tcaaaaatgg  | 6120 |
| aaataggctg | ggtgtggtgg  | ctcctgcctg | taatcccagg  | attttgggag  | gccagtggag  | 6180 |
| gaggattgct | tgagcccagg  | agttcaagac | tagcctgggc  | aacacaggga  | gacctcgtct  | 6240 |
| ctacaaagga | agttttttta  | aagaattagc | cgggcctgat  | ggtggcacat  | acttgtagtt  | 6300 |
| ccggctactt | gggaggctga  | ggcaagaggg | ttgcttgagc  | ccaggggttt  | gaggctgtag  | 6360 |
| tgagccatga | tcagtgtccct | gcactccagc | atgggcaaca  | gagcgagacc  | ctgtctcaaa  | 6420 |
| agaaagcaaa | gggaggggaa  | tacagtatat | cttttgtttt  | ataactacca  | aaattaggaa  | 6480 |
| tacttaccat | ttcttggcta  | aactttatat | tttgattttt  | aaaacttggt  | aaaaattgca  | 6540 |
| atgagaagga | aatttcagga  | gagcagaaga | cagactgtcc  | caggtgtcac  | tgtcctatta  | 6600 |
| ttccctataa | aatccagtg   | caggatggat | gaatggataa  | agcaaagtgt  | gtataagttg  | 6660 |
| aatatccctt | atctaaaatg  | ctttggacga | ggagtgtttc  | gaattttaga  | atatttgcac  | 6720 |
| tatactaacc | agtttaagcat | ccctaactct | gaaatccaaa  | atgctctagt  | gaacattttc  | 6780 |
| cttcagcatc | catcatgttg  | gcactcaaaa | agtttttagat | tttgagtggt  | tttggatttc  | 6840 |
| agattagggg | tactcagcct  | gtgtttgggg | gtagccatct  | cttcatatag  | acatttcaga  | 6900 |
| acttaaatat | tgctttgcta  | taattttctg | gaatttttga  | tatattatct  | tctctgagct  | 6960 |
| acatttttat | cctttataaa  | atggccatat | tgaagtgatg  | atctatccta  | atctaccatg  | 7020 |
| gctgagctaa | gggataaaga  | ggtttttctg | tgtctgtggg  | gtataactta  | cttggtgggt  | 7080 |
| tttatctaga | agcttggttt  | ggtcaagatg | ttggttatat  | tcaggccagg  | catggtggct  | 7140 |
| catgcctata | atctcaacat  | tttgggaggg | caaggtggga  | ggatcacttg  | agctcaggag  | 7200 |
| tttgaaacta | gccagggcaa  | catggcaaga | ctccatctcc  | aaaatttaaa  | ttaaaaaaag  | 7260 |
| atactatctg | tattcatagt  | tgtgtctctt | ttgcctttag  | tccaagctca  | ccttaacccc  | 7320 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| atgggtcggc  | cttcggaaga  | tcaatgtgtc  | ctactggtgc  | tgggaagata  | tgtccccatt  | 7380  |
| tacaaatagt  | ttactacagt  | ggatgccgtc  | tgagcccagt  | gatgctggat  | tctgtggaat  | 7440  |
| tttatcagaa  | cccagtactc  | ggggactgaa  | ggctgcaacc  | tgcatcaacc  | cactcaatgg  | 7500  |
| tagtgtctgt  | gaaaggcctg  | gtaagttcac  | aggtgaatta  | ggtgggtattc | agagtttatt  | 7560  |
| gtgagagaaa  | ccataggagg  | catagtccat  | tgctgagatg  | tgtgaagtag  | tcatgaaaac  | 7620  |
| agatgaagta  | ttgattttcaa | gcatgcaaag  | aagagtataa  | cccagatttc  | agaagcagaa  | 7680  |
| ggaaatatctc | tgggaccctg  | aatagtttta  | attataagca  | aaactaaaaa  | taactaacac  | 7740  |
| tactcgaaga  | aactgatatt  | ctaattaaca  | atgagattga  | taggtttatt  | gaccagaaaa  | 7800  |
| agtattgaga  | attgctctga  | aaagcaaatt  | tattggtgtt  | ggcagagaaa  | tgctgtggaa  | 7860  |
| gaagaaacaa  | aaaaggaaaa  | taaaaccaa   | gaaaagatta  | gtaaagcaag  | caagtgactg  | 7920  |
| cagggacagt  | gttcagaaaag | gtagtgtcaa  | cagggagaaa  | aatgatgaga  | gcagttctgt  | 7980  |
| aagcgaggac  | agaagcaacc  | cagaactacg  | cgagagctgc  | aggagagtag  | tgagcagaca  | 8040  |
| gacactcgga  | gtagcttccc  | agcttttcagt | ctctctgggc  | tacatttttg  | ctaactaaag  | 8100  |
| gcaggcccag  | gggtagctgc  | tgcatcatga  | agacaactca  | aaataagcct  | cctcctcctg  | 8160  |
| cagagggact  | cacaagcagt  | cggtggaatt  | gtgggttatt  | ttgggagggt  | ggactccatt  | 8220  |
| tgcatgtgtg  | ccaatttttg  | gaaagtaatt  | tgtctgagga  | attacagagg  | tataggagag  | 8280  |
| aattcatagc  | attgtagatt  | tctaaatatt  | gatttctaaa  | cttctctagg  | ggagctgccc  | 8340  |
| agcagccttt  | cagaattctg  | aatcctttgt  | tgaacttaat  | gaatgccgta  | gaccttcttc  | 8400  |
| cctccaaaat  | tgcaaacatg  | agattttaca  | tattcaagtg  | gattgtaaaa  | cccctaaagt  | 8460  |
| tcactcctggg | tccccaaatt  | ctaaggggct  | tacagcccta  | cttttatgga  | aagattcact  | 8520  |
| gtttcctcat  | ctttgtcctt  | aatgctcctt  | gaaaagaaaa  | tttttttttt  | ccacaagtgt  | 8580  |
| gtaaatacac  | ttgactaaac  | agtacttgat  | gacttttatt  | gtttcttatt  | ttctcctatt  | 8640  |
| taattgtattt | catggcacat  | taatgggtca  | tcaccacat   | ttgaaaagtc  | ttggctgggc  | 8700  |
| acagtggctc  | acgcctgtaa  | tcccagcact  | ttaggaggcc  | aagacgggtg  | gatcacaaag  | 8760  |
| tcaagatata  | aagaccatcc  | tggctaacat  | agtgaacccc  | catctctact  | aaaaatacaa  | 8820  |
| aaaattagcc  | agacgtgatg  | gcgcaccctc  | gtagtcccag  | ctacttggga  | ggctgaggca  | 8880  |
| ggagaatcac  | ttgaaccggg  | gaggcggagg  | ttgcagttag  | ccgaaattgt  | gccactgcac  | 8940  |
| tctagcctgg  | agacagagtg  | agactccgtc  | tcaaaaaaaa  | aaaagaaaag  | aaagaaaagaa | 9000  |
| aagtctttga  | ccttagcgga  | catggtggga  | gctctaagtg  | tctctcttgg  | gtttcattcc  | 9060  |
| cagcaaacca  | cagtgtcaag  | cagtgcggga  | caccatgtgc  | cttgaggaca  | gcatgtggag  | 9120  |
| attgcaccag  | cggcagctct  | gagtgcattg  | ggtgcagcaa  | catgaagcag  | tgtgtggagt  | 9180  |
| ccaatgccta  | cgtggcctcc  | ttcccttttg  | gccagtgtat  | ggaatgggat  | acgatgagca  | 9240  |
| cctgcccccg  | taagtgaaaa  | aggagaccct  | aggcacttat  | gcatgccctc  | tgtataggca  | 9300  |
| acaactcagc  | catgaggctg  | tgctgtcagc  | ctctgaacat  | tttagaaaca  | agactggaca  | 9360  |
| tgacctctgc  | tcaaacctga  | ccagagactg  | ccatcgagac  | cttgctgcct  | attgagaacc  | 9420  |
| ttcatacaga  | atcaggcaca  | ttgacagtaa  | ataaatgtaa  | gatagatcac  | agagtacaga  | 9480  |
| aataaactgt  | ccaacttcag  | tgtcatattg  | ctcaatccat  | gtaatatctc  | catatctgaa  | 9540  |
| tttcctaatt  | tgtaaagtaa  | atgctttcca  | atagataatc  | tctaaggtec  | cttttgcctt  | 9600  |
| caacatcctg  | ggattgagag  | aggagggaa   | ggtcatctct  | gttatgtatt  | gggcaaaaata | 9660  |
| ctgggctctt  | tacattcatt  | atctctttta  | aataatcaag  | acagaataat  | atttttgact  | 9720  |
| caagccagtt  | gaatagtctg  | ttaaaaaaaa  | gtgaattcag  | atctacttgt  | actgtacctg  | 9780  |
| gatagtcaat  | tgcaactttt  | tttttttaat  | agctgaaaat  | tgttcagggt  | actgtacctg  | 9840  |
| tagtcattgc  | ttggagcaac  | caggctgtgg  | ctgggtgtact | gatcccagca  | atactggcaa  | 9900  |
| agggaaatgc  | atagaggggt  | cctataaaag  | accagtgaag  | atgccttcgc  | aagcccctac  | 9960  |
| aggaaatttc  | tatccacagc  | ccctgctcaa  | ttccagcatg  | tgtctagagg  | acagcagata  | 10020 |
| caactggtct  | ttcattcact  | gtccaggtaa  | gatgccttgc  | atatccaaat  | tcaagtgttt  | 10080 |
| cactactgat  | ttatgaagaa  | taaaaccttg  | aaagctacgt  | tgtgtatatg  | taactccctg  | 10140 |
| ccctcagccc  | ctttccttcc  | tctaatgggt  | tggtaacaag  | aggaatagac  | cagaagctgg  | 10200 |
| tccaaggcct  | gacctggacc  | tgctgagagt  | ggtgggtgggt | tcctaagaaa  | ccaattctaa  | 10260 |
| gaaattggcc  | tttgattcag  | acttgaagtg  | accactcagc  | aatgtgtctg  | tgggtttcta  | 10320 |
| gaacagttgg  | gagaggctgg  | gctgggtcaa  | agactcctca  | gagattagca  | gtcaagaact  | 10380 |
| tctctaagag  | cctgccattg  | acaacagggc  | tgtttgtgag  | gactttgtaa  | gggaaagtcc  | 10440 |
| actgtaaaca  | aagctaaaag  | ggcagagaca  | gactgggaga  | aaatacctgc  | actgcatgta  | 10500 |
| acaactgatg  | atcatccaga  | atatgtaaac  | tcccaccctc  | cagggaaaaa  | ataagaagct  | 10560 |
| aaatgtgaac  | ccaacagaaa  | aagtgggttat | tggagataaa  | gaagccattc  | tcagaaaagg  | 10620 |
| aaatagaaca  | ggaaaatgta  | aactaacacc  | aggaaccaat  | tttttgtcta  | ctaaactgga  | 10680 |
| tcaaattttc  | cgtgttccct  | tttttccata  | ctaagatatg  | ggggacctgc  | aattctattt  | 10740 |
| ttaatctgct  | aggagttaac  | tttttaacga  | aataattaaa  | tctctgcttt  | ttcatgaata  | 10800 |

10/154

|             |             |             |             |            |            |       |
|-------------|-------------|-------------|-------------|------------|------------|-------|
| tcacgaatat  | atctggtataa | atgacaaccc  | agagaatggg  | agaaaatatt | tgcaaagtat | 10860 |
| atatctaata  | agaatccaat  | gtccagaaca  | cgtaactcct  | aaaactcaac | aatagaaaga | 10920 |
| caacccaatt  | aagaaatgga  | taaaggatttt | aaatagacat  | atgtccagag | aagaaatata | 10980 |
| aatggccaat  | aagcacatga  | aaagatactc  | aatatcattc  | atcattacca | agaaatgcaa | 11040 |
| gtcaaagcca  | caatgagata  | ccactttaca  | cccactgaga  | tggctgtaat | caataaaaca | 11100 |
| ggtaataaca  | agtattggaa  | ggatatgtag  | aaattggaac  | tctcatgcag | gttggcaggt | 11160 |
| cctcaaaaag  | ttaaataatag | agtattcata  | cggcccagca  | gttttactcc | taggtacata | 11220 |
| cccaagaaaa  | ttgaaaacat  | atgtttacca  | aaaacttgta  | tataaatgct | tatgottata | 11280 |
| gcatcattct  | tcatagcatc  | atgctcaaag  | caacattatt  | cataataagc | aaaagtgaag | 11340 |
| acaagccaaa  | tacctgttag  | ctggcaaatt  | gatgaacaaa  | gtgttgata  | ttcatacagt | 11400 |
| gtaattgtat  | ttggcaataa  | aaaggaatga  | agtactgaca  | tattttctga | catggattac | 11460 |
| cctaaaaaaa  | catgctatgt  | aaaagaagcc  | aggtgcaaaa  | gattatgcgt | tgcatgatgc | 11520 |
| catttatatg  | aaatgtccaa  | agacagaaa   | tagatgttta  | gtggtttcct | agggctgggc | 11580 |
| atggtaatga  | agaataatag  | gcatgaggtt  | ttctagagta  | gctgcagcat | tattttctga | 11640 |
| catggattac  | cctaaaaaaa  | atgctatgtg  | aaagaagcca  | ggtacaaaag | attatgcatt | 11700 |
| gcatgatgcc  | atttatatga  | aatgtccaaa  | ggcagaaaag  | agatgtttag | tgggttccta | 11760 |
| gggctgggga  | tggaaacaag  | gaataatagg  | catgaggttt  | tctagagtag | ctgcagtatt | 11820 |
| attttctgac  | atgaattacc  | ctaaaaaaac  | atgctatgta  | aaagaagcca | ggtacaaaag | 11880 |
| attatgcatt  | gcatgatgcc  | atttatatga  | aatgtccaaa  | ggcagaaaag | agatgtttag | 11940 |
| tggtttccta  | gggctgggga  | tggaaacaag  | gaataatagg  | catgagattt | tgtagagtag | 12000 |
| ctgcaccatt  | ttatgctctc  | accaacatca  | tatgaaggtt  | catatccttc | cacatccttg | 12060 |
| ccaataactg  | ttactatctt  | tttaatgaaa  | actgttctag  | tggatgagaa | atggcatctc | 12120 |
| actgttggtt  | tcatttgtat  | tttcttgata  | actaatgaaa  | ttgaacatca | acttcattag | 12180 |
| ttagcctttt  | gtatatcttc  | tttggagaaa  | tatttagtca  | aattcctttg | ccatttttca | 12240 |
| gttatgttgt  | cttttttatt  | attgagttgt  | aagagttctt  | tatgaattct | gaagtcctct | 12300 |
| attggatata  | ttatttgcaa  | atactttctc  | ccatttatatg | ggtttctttt | cactttctta | 12360 |
| atatgccttt  | tgaatcaaaa  | agttttcagt  | tttgattaaag | ttccatgtat | cagtgtttta | 12420 |
| ttttatccca  | tgtgcttttt  | ggtattgtat  | ctagaaaatc  | agtgcceaag | taaccagga  | 12480 |
| tcacatagat  | ttgctcctaa  | gttttctttg  | aaaagtttca  | tagatttgtg | tgttacatta | 12540 |
| ggtccttgat  | ccattttgag  | ttaatttttg  | tgtatgggtg  | gaggtagggg | tccaaactct | 12600 |
| ctttttgcct  | gtggatgtgg  | tcctgcacca  | tttgttgaaa  | agattttttt | ttttttttaa | 12660 |
| ccattgaatt  | ttcatggcac  | atttgttgca  | gatcagttga  | ctgtgaatct | aggaactttt | 12720 |
| ttctaagctc  | tcatttttgt  | gcagttgacc  | aatgtttctc  | cttttgccaa | tattacactg | 12780 |
| ttttgattac  | tgtggctttg  | tataagtttt  | aaaattggga  | aggctaagtc | ctccaccttt | 12840 |
| gttcttattt  | tgcaagactg  | ttacagctat  | tctgggtttc  | ttgcatgttc | atattaattt | 12900 |
| taggatgagc  | ttgccatttt  | ctgcaacaac  | aaaaagccaa  | ctggtttgat | aaggtttgca | 12960 |
| ttgaatctac  | tgaccaattt  | ggggtgtatt  | gcctgtttgt  | ttgtttgttt | gtttgtttgt | 13020 |
| ttgtttcttg  | agagagggtc  | tcactctgtc  | accaggtctg  | gagcgcagtg | gtgctatctt | 13080 |
| ggctcactgc  | agcctccgct  | ttcccaggct  | caagtgtatc  | tccagcctca | gcctctcgag | 13140 |
| taactggtag  | tcacagcatg  | agctaccaac  | accagctaa   | tttttatatt | tttttagaga | 13200 |
| cagggtttcg  | ccatgttgcc  | caggctggtc  | ttgaactcct  | gagctcaaa  | caatcctcct | 13260 |
| ggcctttgoc  | tccttagttc  | tgggattaca  | ggtgtgagcc  | accacgcctg | ggcccttgc  | 13320 |
| cattttaaca  | agggttaatc  | ttctagcctg  | tgaacctgga  | atgtcaattt | attacctctt | 13380 |
| ttactttctt  | ttggcattgt  | tttatagttt  | tcagtgtaca  | tcattgtctt | gttctttttt | 13440 |
| ttaggggaaa  | ataattcagt  | ctttattaaa  | tatgatattt  | gtttagtttt | ttttatacat | 13500 |
| gccctttatt  | gggttgagaa  | agtttccctc  | tattctagtt  | ttttgagggg | tttgtaatga | 13560 |
| atgagtgtca  | gattttgtca  | aacacttttt  | ctgtgtctac  | ttagatgata | acgaagtttt | 13620 |
| gttcttttatt | aatactatgt  | attacattaa  | ttaagaatgt  | taaaccagtt | ttgcattcct | 13680 |
| gggataaagc  | ccatttagtc  | atgatgtata  | cttttttatc  | atgatgtata | tttttaactc | 13740 |
| gctgttggtg  | tctatttgct  | agttggggat  | tttgtgtgta  | cattcatgag | ggggatataa | 13800 |
| gtttgtgggt  | ttcttctttt  | gtgattacta  | tctgggtttc  | atagcagggg | agtactggcc | 13860 |
| tcctagagta  | agttggaacg  | tattcctcct  | cctctattta  | ctggaagagt | tcgtcaagga | 13920 |
| ttagcattaa  | ttcttcttta  | gatgttgatt  | gaattcacca  | gtgaaaccaa | atgggcctgg | 13980 |
| cctttttctt  | tacaggaaaa  | tttttaatta  | ttaattcaat  | ctgtttgtta | tagatctatt | 14040 |
| cagattttct  | ttttctgagt  | gagtcagttt  | tggtaactctg | tttttctagg | aatttgtcta | 14100 |
| ttctcagtaa  | tcaaatttgt  | taacatacag  | ttcgtagtat  | tttcttttgt | tttttttttt | 14160 |
| ttttttttga  | gacagagtct  | cactctttca  | cccagctctg  | agtgcagtga | cgtgatctct | 14220 |
| gcttactgca  | acctccgcct  | ccagggttca  | agtgattctc  | ctgcctcage | cccccaagta | 14280 |

11/154

|            |             |             |             |             |             |       |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| gctgggacta | caggcgacag  | ccaccactcc  | tagctaattt  | tttattttgt  | tttgttttgc  | 14340 |
| tttttttagt | agagacaggg  | ttttaccatg  | tagaccagga  | tggctctgat  | ctcctgacct  | 14400 |
| cgtgatctgc | ctgccttggc  | ctcccaaagt  | gctgggatta  | caggcgtgag  | ccaccacgcc  | 14460 |
| cgccagttc  | atagtatttt  | cttgcaaggt  | gctcttttagc | aatatgtagt  | ctcctgttta  | 14520 |
| acctgttaac | gatttttttt  | tctgtctaat  | ttggttttat  | tattctttta  | cattttgcaa  | 14580 |
| gggctctaaa | tattgtgagg  | gttttttttt  | ttaagagctt  | gctctcttat  | attgtggatg  | 14640 |
| caataattta | agccttattg  | aatgtactaa  | aattcgtatt  | tttcattctc  | tcataattct  | 14700 |
| tgcattaacc | cttccctcag  | gatttggtgc  | tgtgcgtggt  | tatcctgggtg | gtattcggag  | 14760 |
| attgagagtt | ttcttatact  | ggttaatcct  | caagggttaat | ttgtattaat  | aactaaatgt  | 14820 |
| gtagatctgt | caatattggg  | cagtgggttg  | gtgtgctgag  | ctggtgtgaa  | agtgtgggtat | 14880 |
| gatgtcctgg | aaaagctagc  | aagccagcct  | caggacactc  | cagctctggg  | ccacttcgac  | 14940 |
| tgttgttgac | tgggtgtgoc  | atctccctca  | ttggaagaag  | gaaggagcaa  | tggacctggg  | 15000 |
| tggctgggtg | caattagcat  | gttcatgggt  | gtggcaagtg  | cctcccagcc  | ttgggtggta  | 15060 |
| gggccaccta | aagatagctg  | gcacattgcc  | ttgagttttc  | tttcattctt  | gcttgcttgc  | 15120 |
| ttgcttttat | ttttgagatg  | gagtttcact  | cttggtgccc  | aggctggagt  | gcaatgggtg  | 15180 |
| tttcttggct | cactgcaacc  | tctgcctccc  | gggttcaage  | gggtctcctg  | cctcagcctc  | 15240 |
| ccaagtagct | gggattacag  | gcacctgcca  | ccacctctgg  | ctaatttttt  | gtatttttag  | 15300 |
| tagagacagg | gtttcactgt  | gttggccagg  | ctgatctcga  | actcctgacc  | tcagggtgatc | 15360 |
| cacctacctt | ggcctcccaa  | agtgatggga  | ttacaggcat  | gagccaccac  | accagctga   | 15420 |
| gttttctttc | aaacactcaa  | atactaacag  | gtgcataaac  | agaacaagta  | caaggctatg  | 15480 |
| gaaggttgcc | aagattggtga | aaaactgcat  | cacactgatc  | aaccaaggca  | gatctgaaga  | 15540 |
| atcagatttt | gcttacagtg  | cagggtgtgc  | actatgtaaa  | tgtttccccc  | atcctattcc  | 15600 |
| cccttagagt | tctgcagtg   | acagcctgtg  | cagctgtcct  | gagcagtcct  | cagaggtcca  | 15660 |
| agcttcatat | acctttctag  | ttgtaaagcc  | ctatatccat  | taggagtcct  | agtaatccca  | 15720 |
| cttgtggaaa | tgtttatccg  | taaagatggg  | tatttttcta  | ataatgaaaa  | atgtaaagca  | 15780 |
| aactaaatat | tagaatataa  | gggaatgggt  | aaataactgt  | taagattcat  | caaaatcatt  | 15840 |
| ctaccatttg | aagttttggt  | tataattgat  | gaaataagg   | ggtaaaatag  | ttgtgatcta  | 15900 |
| agttaagaaa | ggtgaatat   | aatatgaaca  | taattgtgca  | aattatgttc  | ttctgggtata | 15960 |
| gaaaaatacc | agaagaaaa   | atatcaaatg  | tgaacagtgg  | tctctgagtg  | gttggattgt  | 16020 |
| atatgatctt | tccccatcct  | tctatttttt  | aagtatactg  | tcatgagcat  | acgttatttt  | 16080 |
| ttagtgaata | acaacaaaaa  | atatttttta  | atatcactat  | gggctttcac  | cttgcgtgta  | 16140 |
| tacattttca | attttcctgg  | actagctctt  | ctttgatatt  | ttttatctta  | tttctcaaat  | 16200 |
| tagtttttag | tgggccagtt  | ttgggtgatt  | atcaaagata  | aattaactac  | agagatagtt  | 16260 |
| gcagataaat | cttattgaac  | ttaactacat  | gcagttcttc  | catggtacag  | accccttgaa  | 16320 |
| actctttttc | ttgcagcttg  | ccaatgcaac  | ggccacagta  | aatgcatcaa  | tcagagcatc  | 16380 |
| tgtgagaagt | gtgagaacct  | gaccacaggc  | aagcactgcg  | agacctgcat  | atctggcttc  | 16440 |
| tacggtgate | ccaccaatgg  | agggaaaatg  | cagcgttaagt | caaattgggtc | agggttactc  | 16500 |
| atggcaaate | ggtgtggaac  | acagcacttc  | tatttgactt  | gaatatctag  | gaggaaaaaa  | 16560 |
| gccactttgt | tttgataacc  | acattttctta | ttaaatacata | ctgggtcagaa | gctcagctga  | 16620 |
| gctggaagaa | agaaccaagg  | tttcagtgtg  | gctgggttaa  | gcaacaagcc  | aaaaatgtta  | 16680 |
| gagtcaagcc | tttgaccaca  | ttctgcagtg  | gtataagcaa  | gggcccaaaa  | ccagagaaag  | 16740 |
| tgccaacttc | ccctgttttc  | tttctcccca  | ttgaacaggg  | cacaggtcaa  | tagtcagagg  | 16800 |
| gatcagctca | gacatgggag  | tcttggggct  | catctcactg  | aggcccggac  | aaaagggttc  | 16860 |
| agcagtatat | ggacactaag  | gttgtgggaa  | gggcaagggtg | taagggttag  | gactgggaaa  | 16920 |
| gtcctggggt | ggagtcagag  | ttaggaaagt  | gtcttctgaa  | ttccccctcc  | ttccccctac  | 16980 |
| aggaggcctt | ggtgcctgaa  | gaagacctca  | gcccaaggcc  | tcagataggg  | agcctgggaa  | 17040 |
| taaaggctcc | agggctagga  | atgtgaacat  | gtgaaggagc  | gtgacctgtg  | aggtagacct  | 17100 |
| gagtccttaa | ctgcacctct  | ctcagagccc  | cacacttagg  | cccacctcac  | aggctcatgt  | 17160 |
| ctgccaggta | aagcctgtgg  | cattgcctgt  | ggtcaggccc  | tgaaaaatca  | cacacgggct  | 17220 |
| ttttaaccag | gagccagact  | cctcaatact  | tatagtgtat  | tttaaagggtt | taaaatagtc  | 17280 |
| cgtgttggct | gggtgtagtg  | gtcaccgcct  | gtaatcccag  | cactttggga  | gggccaaggc  | 17340 |
| agggtgatca | cttgagggtta | aggagtttga  | gaccagcctg  | accaacatgg  | tgaaccccca  | 17400 |
| tctctactaa | aaatagaaaa  | attagctggg  | catgggtggcg | ggggcctgta  | atccaagcca  | 17460 |
| ctcaggaggc | tgaggcatga  | gaattgcttg  | aacctgggag  | gcggagggtg  | cagtgaagcta | 17520 |
| agatcacgcc | actgcactcc  | agcctgggca  | acagagttag  | actcagctct  | aaaaaaaaaa  | 17580 |
| aaaaagaaaa | aaaaaaaaagt | cctgtgtttc  | ccattttatg  | tcaaatagaa  | gcctgcaagc  | 17640 |
| aataggacat | tttattatga  | gaacaaaaat  | tatgacaagt  | gatattaatg  | aactgctttt  | 17700 |
| tttattctag | tttccagcca  | aaaaacataa  | tgagtaacta  | agtcgaaaga  | cttttcaaag  | 17760 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ttctgagcag  | gaaaagtaaa  | caacataggg  | aaatctctac  | ctactcccca  | gatcccaact  | 17820 |
| ccaaactcct  | tggcatggcc  | cttcagaatc  | tgatctgac   | caagaccagc  | caggagagag  | 17880 |
| gcagcgttct  | agctgtgtgt  | cccttactct  | ccctgggggt  | cagtttcctc  | ctaataatag  | 17940 |
| gggtgggtgt  | aagatgaaag  | cagagcagca  | catgacacat  | gttccacagg  | atactagtca  | 18000 |
| atacttagta  | gatgtgggca  | attcttatta  | cccttctaata | gccattctcc  | actactocta  | 18060 |
| agtaaaagtt  | cttgctcaga  | ttaaaaaggt  | tttggtgaaa  | atgttttccc  | cactctttgt  | 18120 |
| gaatagactt  | aatccgttag  | acagtaacca  | gagtacctaa  | cagagtgggtg | gagtgtacaa  | 18180 |
| ggcattgtgc  | tgttaccac   | tcaagacaca  | ggtgctttta  | ttatccccaa  | gatcaagtaa  | 18240 |
| actgcccagg  | gtctcacagt  | gagatgtggc  | caagctaaga  | accacgcccc  | gtctgtgctg  | 18300 |
| accttaatta  | cttggctggg  | tgtgcttcca  | gatgccatca  | taccacatc   | agcagctgct  | 18360 |
| atctagaagc  | atcacatttt  | cctctgtgag  | atctacaggg  | cctgagtaca  | atttgccctta | 18420 |
| tttttcagag  | tcctaatacca | cggtagaaaag | cgtgggtgtt  | atagaataact | gcagatggca  | 18480 |
| ccaagtcttt  | gatgttttct  | tcttaaaatt  | gtatagtctt  | taggtaacga  | tagtagccat  | 18540 |
| gatttcttga  | atgcttactg  | tgtttgagac  | atcatataca  | tatcatcaca  | aacaccctta  | 18600 |
| tgagataggt  | actgttggtta | tccccatcat  | atggacaagg  | aagctgaaac  | tcagaaagggt | 18660 |
| taagtagcat  | tcccagagtg  | acagaagagc  | caagtagagg  | acgctgaatt  | caaaccagg   | 18720 |
| cagatggacg  | taaggccttt  | gcttttctact | atacattaca  | caccttcccc  | cagtctcaac  | 18780 |
| agagccaagt  | gtcagacact  | catgattctg  | acttcagcat  | ccattctgtt  | aatctctgca  | 18840 |
| ctattctgtt  | ttatgttttt  | ctttgggtga  | ttccaaagga  | tttattttta  | aagcatacct  | 18900 |
| ctctccagct  | gaatgccttt  | catttattca  | ccagcaaaac  | aggggtataaa | gtgaaaagggt | 18960 |
| gttgacaaaa  | aaggcttttca | ctttttttcaa | ctggagtata  | atattttatca | cggcttgtat  | 19020 |
| tacgttggat  | gataaaaagga | gagatgttga  | gttggccaga  | tgagggagat  | gggaagagct  | 19080 |
| attttctgta  | aaggcaaagg  | aaaagggtgaa | gcaagggtgg  | catgtcctga  | gggcccctaa  | 19140 |
| agcgtgacc   | ttggcattgg  | atcctggttc  | tgagccctgc  | ggtgattcca  | aatgtgacca  | 19200 |
| tcgtggagaa  | ggcacctcag  | ccaagaccag  | ccctccttca  | gaggacagtc  | acctccagag  | 19260 |
| cacctcccta  | ggggcagcag  | ctatgtctgt  | gtccccacac  | ccagcacctg  | gcattctaga  | 19320 |
| gatgcttcac  | aaatatattat | tgaatgaatc  | aaagaatcat  | agcagtgaac  | aagagagtct  | 19380 |
| attgaaaaga  | tcaaaatagt  | cattgcttca  | gaaggcagtg  | ggtaccagca  | ttcactttcc  | 19440 |
| tcctgtcctt  | tgtttgggttc | ctttaatgtt  | acctatacat  | tattatcatt  | agctgaaaaa  | 19500 |
| tcatgaatct  | taccatttga  | atgctcgtac  | tttaatctgt  | ctttcctact  | gaactcaatt  | 19560 |
| cagataaaat  | tgcctgtttt  | ggaaaaagtc  | tccaatagtc  | agaatttttt  | aagccagttg  | 19620 |
| tgacctctgt  | tcctttttct  | tctctgcagc  | atgcaagtgc  | aatgggacag  | cgtctctgtg  | 19680 |
| caacaccaac  | acgggcaagt  | gcttctgcac  | caccaagggc  | gtcaaggggg  | acgagtgcca  | 19740 |
| gctgtgagta  | ccatactccc  | tggaccacca  | gggaggacca  | agaggctgtg  | cagctgcctg  | 19800 |
| aacccccacc  | tgagagccac  | ccacttccct  | gtgtcttgtt  | gctgtgggct  | ctgagggatc  | 19860 |
| cctgggttga  | ttagtttgaa  | attttgcccc  | ttctatttca  | gacaggctcag | ttcccaaaat  | 19920 |
| gaggaggtcg  | tcgagtttagc | agcaattcct  | taatggctct  | tgaattcaca  | tttttgttta  | 19980 |
| aatgatactg  | acatttccctg | ggttgtccat  | ttggagtagt  | catttttaact | tcagcaacta  | 20040 |
| cttgattttt  | gtcatgtcaa  | gagattatac  | tccttcccaa  | agagtagaga  | tgggaagagct | 20100 |
| ggtgttttgg  | tgatggcatt  | tatttggcgt  | ttggttgtct  | gattgtggaa  | atgatcctct  | 20160 |
| tacctttaac  | atttcccat   | actctagcat  | tttcttgtt   | gaagcatgtc  | aggcatcttc  | 20220 |
| ctggagaggg  | tttgaagtcc  | ccatccatga  | ggatgcaggt  | gcagcagcat  | caggcatgtt  | 20280 |
| taagggtttca | gacctggccc  | tgcctccact  | gagcaagggt  | ctcttgaggc  | aaatcactgt  | 20340 |
| tccccccatc  | tgtcagtagt  | gggactagca  | gtaattgttt  | ttggtgtcct  | ggtaggttga  | 20400 |
| aacaggaagg  | atagttcctc  | aatcaatgta  | taccgggctg  | tctcactggg  | aaacctcata  | 20460 |
| aatgcatgt   | atccgcattt  | actcattctc  | aagaaaaaac  | ttaaaaagtg  | tttagtgtca  | 20520 |
| aaacactgag  | ctgattaatc  | attgtaaaca  | tcattatttc  | ttaaataaca  | gtagtaataa  | 20580 |
| tgccaccagt  | gatttgagaca | tcattgctgg  | tgctactcag  | cagggtttgg  | tctgtctact  | 20640 |
| ttttagttta  | tatctgcata  | tatgtttact  | agcttcatgt  | tgcctaaaag  | ccaatatgtg  | 20700 |
| aagaaaagtc  | taatccaagg  | ttaaatttaa  | cttttaggtc  | cacattccct  | taaacagaaa  | 20760 |
| aatgtcatcg  | agcaaaatta  | atcactctta  | ccctcagtta  | gctgatgcaa  | caaagacatg  | 20820 |
| aggtggcatc  | aagtaagact  | ttcaagttcc  | ttcgagatcg  | cagatgtagg  | cagggtgctgc | 20880 |
| tcatcaccct  | ggctaaaggg  | acacagcata  | cctgcctggg  | aaaggccatt  | accacttcc   | 20940 |
| cgccttttcc  | tcttcatgca  | tagctgttgt  | gtttacctaa  | atgcttcatg  | aagtgggagg  | 21000 |
| ctgggttttg  | ctagatgttta | tagatgatct  | gtatggggaa  | attgttttct  | gagtagaaca  | 21060 |
| tatttttttt  | ctacagactg  | gtgaacactt  | gttccagggt  | ttaaaaaac   | agtgtattct  | 21120 |
| ggtatttagt  | cttctctcac  | ttgtgtttta  | agaagagaat  | tagtttatcc  | agggaaacata | 21180 |
| ggaaagaagt  | gagaaattag  | tatttgaaaa  | atgttttgac  | tgcactttta  | gaagaatatg  | 21240 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tagtccacac  | aaaattgaaa  | atgttttttca | ttatagtaaa  | taggttttatt | acatttggat  | 21300 |
| ttctctcatt  | gcttttgcaa  | atgcttgtca  | attgtcctca  | tatagttggt  | gtgtgaactg  | 21360 |
| ctgagcagtc  | agtattgaag  | cgttcatgca  | tgctcttcct  | agactgcctg  | agctgagtta  | 21420 |
| tggtgaaggga | tgcaattata  | atggctccag  | actatctgta  | ctttatataa  | aaggatctgt  | 21480 |
| ttggtttttaa | aatgaattca  | gtttctgttt  | taaatagcag  | tataaaatag  | tcttttttcc  | 21540 |
| ccccagatg   | tgaggtagaa  | aatcgatacc  | aaggaaaccc  | tctcagagga  | acatgttatt  | 21600 |
| gtaagtgggt  | ttgcaattct  | tatttctaga  | agcaaagtag  | ttcagtaaaa  | cttcattgtt  | 21660 |
| taaacgggtt  | tgagaatagt  | aagtgtctata | agactatagc  | agccaccaat  | gaagtgttcc  | 21720 |
| cagacttgat  | atgtttacat  | tctgttaagt  | ttactacata  | taggagcact  | gttttaagct  | 21780 |
| gttttaattg  | tgtttggggt  | taacgttaat  | gtgtccatag  | cagatagcag  | gagagagtag  | 21840 |
| agaggcatgc  | atctttgtct  | atccacattt  | atgttctctt  | aaaactttac  | tttatttggc  | 21900 |
| attacctagt  | tggggtcac   | atatttctgt  | ttttaggatg  | tagatcaaaa  | acagaattct  | 21960 |
| tacaatttgg  | ttgaactttt  | tgtcattctg  | tctggaactg  | atgtgtgta   | accaccttga  | 22020 |
| ggtgaaaggt  | gagctcgaca  | aggtgacct   | gtttttatgg  | gtaaatgtgt  | tttctcttta  | 22080 |
| tgggagaatc  | cacatggtag  | accagagtac  | gggaccagaa  | aaaaggagtt  | aatgttatgg  | 22140 |
| catatccatt  | gtaattatat  | acctgctgta  | ctgggttgaa  | taacatccct  | tcaaaattca  | 22200 |
| tgctccactg  | gaacctgtga  | atatgacttc  | atatatatat  | atataaaact  | tctaaaagaa  | 22260 |
| aacaagagaa  | aaacatcttg  | accttgggtg  | gggaaaagtt  | gttagatagg  | attctaaaag  | 22320 |
| cataattcat  | aaataaaaaag | tagactaatg  | aaactatata  | aaaattttaa  | acttgtagga  | 22380 |
| aaaaaccatc  | caaaacataa  | aaagccacag  | actagagaaa  | atactcgcac  | atcatcatat  | 22440 |
| atlttgataag | gagtttatat  | ccagaatata  | tgaggaactc  | ttacaactga  | ataagatgac  | 22500 |
| aattccagtt  | aaaaacaggc  | aaagatttga  | aaagacattt  | catgaaagaa  | gatattacag  | 22560 |
| tgggcaataa  | gcacatttaa  | aacgacttaa  | catcttcgt   | tatttggat   | ggagatttta  | 22620 |
| aaaccacagt  | cagatacagc  | cctaaagtga  | ttagaatcca  | acgctaattg  | catggctttt  | 22680 |
| tagaagacag  | tggtaatctc  | atgtctgctt  | ctgcattcag  | tctgttgacg  | tacatctttt  | 22740 |
| tggttaaaca  | catgaaaaaa  | acctggcctt  | accaggcatg  | cagttggaaa  | agggatatag  | 22800 |
| gatacccttt  | taatagcctt  | ttcagataat  | tatggacgtt  | acttgatatt  | atgctgaaac  | 22860 |
| tggaacaactg | gtcattttctt | taaagcgggt  | tgatgggtta  | acccaaataa  | aatgatactg  | 22920 |
| cccacagact  | tgattttctt  | tgaggaacct  | gggtacagct  | gagttaaagt  | gctttcctcg  | 22980 |
| ttacttaatt  | tataagaaaa  | gcagcctgta  | tctctaaaga  | cctgttctat  | tttgtgtgtg  | 23040 |
| tttagttttt  | aaatatgcat  | ttctttcttt  | ccatagatac  | tcttcttatt  | gactatcagt  | 23100 |
| tcaccttttag | tctatcccag  | gaagatgctc  | gctatttacac | agctatcaat  | tttggtgcta  | 23160 |
| ctcctgaoga  | agtaagattt  | tttaaagtct  | tctatttttg  | ttttgaattt  | gtatggatct  | 23220 |
| ttttcttgggt | cattacggat  | ggacgtactg  | ccttaacagt  | gctctccaga  | ctggagtaca  | 23280 |
| cgagatgac   | tctagaggta  | taggaaagaa  | atgttagact  | ctacgttatc  | tcttttccac  | 23340 |
| ataaaaaggca | aaagtgatgt  | taataagatt  | taccaggatc  | ttagacacag  | actgacattg  | 23400 |
| attccacgca  | tacttactct  | gcctgtcagc  | ccatcatggc  | ctcatacaga  | aaggggactc  | 23460 |
| caccatcaga  | gggcagatag  | cagagcctgg  | gtttactctc  | tcaagagtga  | ccagaggctt  | 23520 |
| aaagacactg  | ctgattgaaa  | tgccagtgtat | gcagcccaa   | tcagacagca  | agggaggggga | 23580 |
| ccccaatcag  | acagcaaggg  | aggggacccc  | aatcagacag  | caagggaggg  | gacactctgc  | 23640 |
| tcctagagtg  | agttcttagc  | ctcattagga  | ggcaaaacag  | caaaggctta  | gttgggtcca  | 23700 |
| ttaagaagtt  | agccaggaat  | aattttaaatt | gttaaaatat  | gtatgtaaaa  | tgtggatttt  | 23760 |
| tttatctgct  | gtcatttaacg | atgaagccca  | acctgcctta  | aggtattacc  | tagtggtaga  | 23820 |
| aggaaggcca  | cactgcggga  | cattttaaacc | tgaacatac   | agaacacgaa  | gatgcacctg  | 23880 |
| tacagtttct  | tcaatgaaat  | ataaagtcac  | gcagtaccca  | cttcagtatt  | taaagaaatt  | 23940 |
| ttggtaaaaca | taatggtaaa  | ttatttagga  | acttccttgg  | agattttctta | cttctcatga  | 24000 |
| acatacacaa  | agccattttac | cattacaaaa  | ttccatttaac | aataaatgtg  | acaaataata  | 24060 |
| catggaaaac  | aatatggcag  | taagacaaag  | tatattgctt  | tgtttaacaaa | tgttatctat  | 24120 |
| aaacatttgc  | aaattttaat  | ttaaaagttag | aacaaagtct  | atagtgtggt  | atgtttacta  | 24180 |
| tacaagttaa  | tgaaaatagg  | atttgttttt  | aattttctat  | ttcaaagata  | aaattgaagaa | 24240 |
| aaaatgtact  | agaaaagata  | cattctcaaaa | agtaaatgtc  | tttttaagta  | aaaataataa  | 24300 |
| attacttttaa | atgaattatt  | ttatggaaaa  | actataggta  | ttaatatata  | atlttgagtgg | 24360 |
| ctgtattgac  | tggaattttt  | ttcttaacga  | ttttagaata  | agatttttaac | aaaagtacca  | 24420 |
| tatatgaaat  | gtatttactg  | cctcataagc  | aagcgtttga  | agtgtcctaaa | atctttcagg  | 24480 |
| atatacttca  | tgccattaat  | gtcatttaaaa | aaataaatat  | agtagaatct  | ttgtaatact  | 24540 |
| tcttaaccag  | gtgggaagca  | accatgaaag  | tatttggacc  | tttctggatt  | tcccagttta  | 24600 |
| tcttacgaca  | gaatactttac | tagagttatc  | caaatgcata  | tgttctgtcc  | tctataaaaag | 24660 |
| cacaagcatt  | ttaagttttat | tgattctttc  | tgtggaaaga  | catatagttg  | acccttgagc  | 24720 |

|             |             |             |            |             |             |       |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| aatacagggtt | aaaactgtga  | ggatccactt  | acacgtagac | ttttttcaac  | caagcgtgga  | 24780 |
| atgaaaatac  | agtatttgca  | agatgtgaaa  | cctgagtata | cagagggtgg  | acttttcata  | 24840 |
| tgcaagggtt  | ctatgggcag  | actgcggact  | ggagtatgtg | tggatttggg  | catgctcagg  | 24900 |
| gggtcctgga  | accgatgacc  | cacgtatacc  | aaaggatgac | tgtaatgtgt  | agttgtgtgc  | 24960 |
| tgccagcaaa  | tactaaatac  | aaataagtaa  | acacttgagc | tgtcctttca  | agatgaagg   | 25020 |
| gaggtccttat | cagtaaatac  | gaaggtaaat  | gctttgtgag | agaaaacttct | ggtttaata   | 25080 |
| tcacattttta | aaaatagctg  | tttggaatg   | tttccattgt | tgtgattttg  | tgtattataa  | 25140 |
| atgatcaaaa  | caacaccctt  | aaaaatctta  | ttctaacctc | tcaagatctt  | ttaaaaatga  | 25200 |
| ataatttcag  | tacagtgcga  | tgcattctgta | aaagataaaa | atataacatt  | gattagtttg  | 25260 |
| caaaaataat  | tgtttgaccc  | cagttaagag  | atgtactagt | caaatttcag  | tttgactaat  | 25320 |
| tattaatggt  | ataatttacc  | taacatcacc  | aacagtacac | ttcctccact  | ggcttaacag  | 25380 |
| attcctcagc  | aatatcttta  | ttagtcatta  | agaccaagga | tcaaaataaa  | ttaagttaga  | 25440 |
| ttagccctgt  | gaactgctat  | atctctaagt  | ggtagacaag | gttttcaaaa  | ctaagaagcc  | 25500 |
| atactcaatc  | atatttctct  | tgaataata   | tttttagtaa | gagaaaaata  | tttttcaaat  | 25560 |
| catgaatata  | ataaaattat  | tttttaaat   | aagtacattt | cagctctata  | tgcctttttt  | 25620 |
| aaaggctggt  | tccagttttg  | gagaagtttt  | acactgtata | taatctatga  | gtttagaatt  | 25680 |
| atatgggttt  | cagttataaa  | taaatataat  | tttgggcact | tgatcaagtg  | ttttattggt  | 25740 |
| aggagtgaac  | ttcagacatt  | tgaacacagc  | tgccagaga  | atccttagga  | gggtgattcc  | 25800 |
| ccagcacagg  | tacaggatg   | caaatacctg  | tgctctgata | ggggcttttt  | taaaaggcac  | 25860 |
| tttctcatgc  | agttaggcat  | ttgataaagc  | attaaattat | gtatacttta  | atggggaggga | 25920 |
| ataaaatttt  | attggcatat  | attgcttgta  | ttaaggaaaa | ttggttagaa  | attttgctaa  | 25980 |
| ctcttctgtg  | agtttctcta  | aagacatcat  | aaaatgtttg | tgattaagaa  | catttagagg  | 26040 |
| agtaaacttt  | attgctttat  | tttaaaatct  | agaaattgtt | ttaattaat   | ttctaataat  | 26100 |
| ttgtaccgct  | catcagcaaa  | acagggtatt  | ggacatgttc | atcaatgcct  | ccaagaattt  | 26160 |
| caacctcaac  | atcaacctggg | ctgccagttt  | ctcaggtaaa | gacataccta  | gagaagaccc  | 26220 |
| tgcaaatgaa  | ggtgtggtag  | attaagaaat  | gtaatatagg | aattgagaaa  | gcgagctcag  | 26280 |
| gagacagatt  | ggtttgaaac  | ccacccttgc  | cacttactag | ctatgagacc  | ttgagcaagt  | 26340 |
| atctaaatcc  | ctctctaacc  | cttagcatta  | ttttattcat | ctgtaaagtg  | aggataatga  | 26400 |
| tacttaacctc | ttagaattgt  | tgtatagatt  | aaattagggt | atacctaacc  | aggcttgctg  | 26460 |
| tggtgtcagg  | ctcagtggtg  | ggttactacc  | ctagcccacc | accaccattt  | ctgttcttgc  | 26520 |
| tgtggccact  | ggcactacca  | tcattgtcta  | catccgtgct | tcggaagtga  | aaaattcaaa  | 26580 |
| tgattcgttt  | caataaatga  | aaacatttta  | aataaaatga | gatttttagta | ggtacagaga  | 26640 |
| aatgtaactt  | gggaattaca  | tcaagctcta  | aaagcacagc | tcttgcctgc  | tgccttactg  | 26700 |
| tgattcactg  | aagatctact  | gtatagaaaa  | tctaagaaaa | taaaggatga  | aggccaagtg  | 26760 |
| ctgtggctca  | tgcctttaat  | ctcaacactt  | tgggagactg | aggcaggagg  | attgcttgag  | 26820 |
| ctcaggaggt  | taagaccagc  | ctgggcaaca  | tagtgagacc | ctgtctacaa  | aaaaaaaaaa  | 26880 |
| aaaaaaaccc  | agggtgtagt  | actcatacct  | gtaatgcca  | ctactcagga  | ggctgagatg  | 26940 |
| agaggatcac  | ttgagcccag  | gggttgagga  | ttgcactgag | ccatgatcgc  | accattgtac  | 27000 |
| tccagtgtgg  | gcaaacagag  | caagatccca  | tctctttaaa | aaaaaaaaaa  | aaaaaagaaa  | 27060 |
| aaaggataat  | cactacttaa  | cttgataact  | caacaagtag | atatgggttt  | gaaatttgtc  | 27120 |
| cattaaattt  | acttgcaccg  | tgtgtttagg  | caagttactt | aagggtttctg | agccagtttc  | 27180 |
| ctcctgtata  | aagtaggata  | gtaaaaacac  | cgtcctggca | gggcgtgata  | gctcatgcct  | 27240 |
| gtaatccag   | cactgtggga  | agccaagggtg | ggaagatcac | ctgagggtcaa | gagttttgag  | 27300 |
| accagcctgg  | ccaacatggt  | gaaaccctgt  | ctctactaaa | aatacaaaaa  | tcagccaggc  | 27360 |
| gtggaggcac  | gtgcctgtaa  | tcccagctac  | tcaggaggct | gaggtaggag  | aatcgcttga  | 27420 |
| accttgaagg  | tagagggttg  | tgtgagccga  | gatcacgcca | ctgcactcca  | gcttgggtga  | 27480 |
| cagagtga    | ctccatctca  | aaaaaaaaaa  | ataaaaaaac | accttcccaa  | gtagagtgat  | 27540 |
| gtgagaatta  | aatgagataa  | taaatgaagt  | actcaatata | gtgcttgaaa  | tgtggtaaat  | 27600 |
| ggtaactata  | ttttatcatt  | attactatta  | caatactggg | tttttaaaaa  | tcaaaaacac  | 27660 |
| aaagcaatga  | gattgatgca  | aaataagaat  | attgccttgt | gcacgccaact | tacgtttatc  | 27720 |
| atcttaaaac  | attgtgtaga  | atttgagaaa  | agttcagaaa | ctctcaatga  | ggagggactt  | 27780 |
| ttaagaaaaa  | gtctgaatta  | tcagagtatt  | tggagaaagg | caacatctcc  | aggcatgtga  | 27840 |
| aagatttgca  | atgagccggg  | cgggtggctca | tgcctgtaat | cctagcaact  | tgggaggctg  | 27900 |
| gggcagggtg  | attacctgag  | gtcaggagtt  | caagaccagc | ctgaccaaca  | tggtgaaatc  | 27960 |
| ccgtccctac  | taaaaataca  | aaactagctg  | ggtgtgggtg | ggcgtgcctg  | taatccagc   | 28020 |
| tacttgggag  | gctgaggcag  | gagaattgct  | tgaactctgg | aggcggagg   | tgcagtgacc  | 28080 |
| caagattgca  | tcatgtgact  | ccagcctggg  | caacaaaagc | gaaattccat  | ctcaaaaaaa  | 28140 |
| aaaaaaaaaa  | aaaaaagatt  | tgccatgagt  | gtctcaatga | agacgtgata  | atgtgggctc  | 28200 |



15/154

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tagtcacagg  | gtctaactca  | gacatggaaa  | aaagtccatt  | tcattaatct  | ttatcggcac  | 28260 |
| ttgaattcct  | ggctaaggga  | gaatgtggaa  | cattgaaggga | ctctctggga  | ataggatgga  | 28320 |
| gttataccag  | attaggggga  | cttaaatact  | gtggtagctg  | gtggtagaag  | ggaggactga  | 28380 |
| gtgacccctt  | gaacccctcc  | tccctgctac  | agtgggttag  | gcagtgagcg  | gtacatcagc  | 28440 |
| attactggca  | tgggagtctg  | gcgcattgcc  | aaggaggtgt  | aaaggggaaa  | tgcaaaggaa  | 28500 |
| ttgaagtggg  | gtgggcaaa   | tgaatgccag  | tgcttgtaa   | taggattcta  | gtggtatctg  | 28560 |
| tattttcatg  | atcatgtgtg  | tcacctgttt  | gggggtgggg  | caaggggtga  | agggagttac  | 28620 |
| atggattcct  | ggtaaaacca  | tttttctttc  | tttctttttt  | tttttttgag  | acggaatttc  | 28680 |
| gctcttattg  | cctaggctgg  | agtgcaatgg  | cgcaatcttg  | gctcactgca  | acctctgctt  | 28740 |
| cccaggttca  | agcgattctc  | ctgcctcagc  | ctgctgagta  | gctgggatta  | caggcatgcg  | 28800 |
| ccaccacgcc  | tgggataatt  | tgtattttta  | gtagagacgg  | ggtttctcca  | tgtttgtcag  | 28860 |
| gctgggtctca | aactccgacc  | tcaggtgatc  | caccgccttc  | ggctctccaa  | agtgtctggga | 28920 |
| ttacagatgt  | gagccaccgc  | accggccag   | taaaaccatt  | ttggttaggg  | gcataggctt  | 28980 |
| gtatatcagc  | ctgcccagct  | ttaaatctta  | tctccatttc  | ttgttgctg   | tgtgacttga  | 29040 |
| gggaagttac  | ttaatctctg  | tgaacctcaa  | tttcccagtc  | tataaatgaa  | gataataata  | 29100 |
| gggcagctgt  | gaggattaaa  | tgagattgag  | ctcttaaagg  | ctattactgg  | aacacaggaa  | 29160 |
| atgtttgata  | aatgctattg  | tccttattat  | taatgaggcc  | agattctgtt  | ctccccctag  | 29220 |
| cccccaaaa   | aatgtctctt  | ctctttcatg  | tttttctttt  | aacagctgga  | accagagctg  | 29280 |
| gagaagagat  | gcctgttgtt  | tcaaaaacca  | acattaagga  | gtacaaagat  | agtttctcta  | 29340 |
| ttcagaagtt  | tgatcttcgc  | aaccacccaa  | atatcacttt  | ctttgtttat  | gtcagtaatt  | 29400 |
| tcacctggcc  | catcaaaatt  | caggtaagaa  | gaggcttttg  | gtctcatacc  | tgcaaagggtg | 29460 |
| gtgaaatctc  | tttagtaaga  | ctaaatttac  | taatttggag  | cttgtggtaa  | atgagatgtg  | 29520 |
| caatgtggct  | ttgcctttgt  | aacgtgtatg  | gcagaggagt  | gctgagcaca  | tgcatgctgc  | 29580 |
| acagaagatt  | ggagtgggga  | tggactgtat  | cactcatgaa  | agacatttgc  | aaaagcactg  | 29640 |
| ttgaaagcaa  | gttggcatgt  | aacagatttg  | gtcattataa  | acgtattcac  | ttcttcagtg  | 29700 |
| agcatttgcc  | atgtgaaagc  | ctgtagggtc  | acacaaagaa  | cttcaattct  | agagtaagggt | 29760 |
| ggatgtaagt  | gaaacaaaac  | cacatataac  | aactgaaagc  | cagagtgtgg  | aaagtgcacat | 29820 |
| gagatgatcc  | cagaaatgct  | atcaaagttt  | aaaggaggac  | aaatggggag  | actatgttga  | 29880 |
| agaacatcag  | ccttctagtc  | agacagaggt  | ggattgattc  | ctggctccta  | tacaaatcat  | 29940 |
| acagcctttc  | caagtctcca  | gttctctgtg  | catgtgacct  | gaagggtagt  | tgtaaagggc  | 30000 |
| tgtgcagctg  | ctgcagtgtc  | ttgttagcct  | gctcctttcc  | tctgttccca  | ggggggcagt  | 30060 |
| gtactccctc  | ttgtccgaga  | cccatggccc  | cattttaact  | ttttatactc  | atgtcccctg  | 30120 |
| gggcctttcc  | tcaatacctt  | ctgcttctta  | ccttcttcat  | ttaggtgaat  | gtggagggtta | 30180 |
| gggatagggtg | ggctttcaag  | gactggttca  | cctttaacca  | tggaaagcatg | gtcactggac  | 30240 |
| ggaggctggt  | gctgtttgcc  | aatgttcaga  | agcataatca  | acctcagaag  | caagtcacca  | 30300 |
| caaacatatg  | aaaaaagttc  | aacatcactg  | atcattagag  | aatgcaaat   | cagaaccaca  | 30360 |
| gtgagatacc  | atctcacacc  | agtcagaatg  | tctgttatta  | aaaagtcaaa  | aaagaacaga  | 30420 |
| tgtctggcaag | gctgtggaga  | aacaggaacg  | cttttacact  | gttgggtggga | gtgtaaatta  | 30480 |
| gttcaaccat  | tgtggaagac  | agtcctggcaa | ttcctcaaa   | acctagaggc  | agaaatacca  | 30540 |
| tttgaccag   | ccatccactt  | actgggtgta  | tatccaaaag  | aatataaatc  | attctgtaac  | 30600 |
| aaagatacat  | gcacatgtat  | gttcattgca  | gcactattca  | cagtagcaaa  | gacatagaat  | 30660 |
| caacctaaat  | gcccacagct  | gatagactgg  | ataaagaaaa  | tgtggtacat  | atacaccatg  | 30720 |
| gaatactatg  | cagccataaa  | aagggaatgag | atcatgtcct  | ttgcagggac  | atggatggag  | 30780 |
| ctggaggcca  | ttatcctcag  | caaactaatg  | caggaacgga  | aaaccaggta  | ccacatgttc  | 30840 |
| tcactcgtaa  | gtgggagccg  | agaacacatg  | gacacatggt  | ggggaacaac  | acacactggg  | 30900 |
| gcctattgga  | gagtgatggg  | gaggaaggag  | agcatcagga  | agaatagcta  | gtggatgctg  | 30960 |
| ggcttaatac  | ctagggtgatg | agatgatgtg  | tgtagcaaac  | cactgtggca  | caogtttacc  | 31020 |
| tatgtaacaa  | acctatacat  | cctgcacatg  | taccctgaa   | cttaaaaatt  | aacaataaca  | 31080 |
| aaaaaagcaa  | gttatcactc  | atcaattagg  | atgccttggg  | actgtgacta  | aagaacagtt  | 31140 |
| gatctttcca  | ttctggaaat  | atggaggaca  | aaaactatgt  | tgtagttttt  | ctcaccaccc  | 31200 |
| tctctccttt  | ttctgtctc   | catggtctaa  | gatttgtaa   | tcctctcatc  | aggttctccc  | 31260 |
| tttgccctcg  | catactccct  | gtgtccctcc  | tcagccagtc  | ttgtaagcat  | cagccacgct  | 31320 |
| tcttattcct  | gttttctctt  | gtctagtcc   | acatctgctc  | acaatgggtc  | ggcctcttcc  | 31380 |
| ctcaccacat  | cctgaaactg  | tcctgtcata  | gtgggcccga  | gtgatggaa   | ttttcagttc  | 31440 |
| cctttttatc  | tcattgcatg  | gctttgaatc  | atcatctttc  | ttgtcttctt  | ggatttttct  | 31500 |
| tttctctggt  | ttctggccag  | tcttttttagg | gaatcctctt  | cctcttctct  | aaacactggg  | 31560 |
| cttgtctagg  | attcagacct  | ggtccttttc  | tctcttcact  | ctgtttctctg | catgtgggac  | 31620 |
| ttgtgggtgcc | acatctattc  | tgggtgattct | cagtgtctgc  | tccagaccgg  | cactggcaggt | 31680 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tgctgcctgg  | acagccctac  | ctgggtagtc  | aaagctcccc  | gcgttcggtg  | catattcctt  | 31740 |
| cttgtgatct  | tagccccaga  | cttccatgct  | tcctttctct  | gtaaacagca  | ccaccatcac  | 31800 |
| cctggtgtac  | aagctagggc  | ctgatgatca  | tttgcattcc  | tgggataaac  | ccaacttgag  | 31860 |
| catgacgtat  | tatttcctta  | ctttacacgt  | tgttggtatc  | agtttgctaa  | tatttagggt  | 31920 |
| agaattttgt  | atctatttca  | tgagtaagaa  | tgtctgataa  | ttttcccttc  | ctgtccttgt  | 31980 |
| tatgggtttg  | atatcatggt  | aaactaactt  | ttaataatta  | taatttgagg  | agttttcact  | 32040 |
| ctttctcatc  | tctggagatg  | taagattaga  | gttaactgaa  | cctcaagtac  | ttggtagctc  | 32100 |
| tctcctgtaa  | aatcatttga  | tctagtgtc   | ttatttgtgg  | agatactttt  | tagctgctga  | 32160 |
| ttcagcttct  | ttaatagtta  | taggatattt  | tgaatttcc   | ctttatttga  | ttcacttttt  | 32220 |
| aatatttttt  | ctaggttttt  | atatcatggt  | tttaaagtga  | tttttaattt  | tactttttta  | 32280 |
| attgatacat  | aattgtacgt  | atztatgagg  | tacatgagat  | attttgatac  | atgcatacaa  | 32340 |
| tgtataatga  | tcaaataaga  | gtaattagga  | tatccatcac  | ctcaaacatt  | tatcatttct  | 32400 |
| ttgtgttgag  | tacatttcac  | atcttctagc  | tattttgaaa  | taatgaataa  | gttattgcta  | 32460 |
| actatagtca  | taatgactat  | tgtgccattg  | aacactagaa  | cttattcctt  | gtaactctat  | 32520 |
| ttttataccc  | attaacctct  | atctccccag  | ccccccagca  | gctggtaacc  | accattctgc  | 32580 |
| actctacctc  | catgagatca  | gttttttttag | ctcccacatc  | tgagtgagaa  | aacatatcta  | 32640 |
| tctttctgtg  | cctggcttgt  | ttcacttaac  | ataatgacct  | ccagttccat  | tcatgttget  | 32700 |
| gcaaaggaca  | ggattccatt  | cttttttgtga | ctgaataata  | tttcattgta  | taatatatac  | 32760 |
| aacattttat  | ttgtccattc  | atttgttgat  | ggacacgtag  | gttgattcca  | tgtcttcact  | 32820 |
| cttgtgaata  | gtgctgtgat  | aaacatttag  | agcatgcagt  | atctctccag  | tatactgatt  | 32880 |
| ttctttcttt  | tggatatata  | cccagcagcg  | ggattgctgg  | atcatgtggg  | ggatctatta  | 32940 |
| tgagcagtc   | tccatactgtc | ttccatagtg  | gctgtactaa  | ataatttaca  | ttcccaccag  | 33000 |
| cagtgaacta  | gtatcgtctt  | tctctgtatc  | ctcgccagca  | tctgttattt  | tgtcttttta  | 33060 |
| ataatagcca  | ttttaactgg  | gatgagatga  | tttctcatta  | tgggtttgat  | ttgcatttgc  | 33120 |
| ctgatggtta  | ctgatgttga  | gattttttttc | atatgcctgt  | tggccatttg  | tatgtcttct  | 33180 |
| ttggacaaat  | ctctattcag  | atcatttgcc  | cattttttaa  | tcaagttttt  | ttcctattga  | 33240 |
| gttgtttgaa  | ttgctgggtat | attctgggtta | taaattccctt | gttggatgga  | tagtttgaac  | 33300 |
| atattttctc  | tcattctata  | agctgtctct  | tcactctctt  | gtttcccttg  | ctttgtagag  | 33360 |
| ctttttggct  | tgatataate  | ccatttgtcg  | atttttgctt  | ttgttgctg   | tgatttcgac  | 33420 |
| atcttacaca  | aaacatcttt  | gccagacca   | atgtcctgaa  | ggattttccc  | aatattttct  | 33480 |
| tctagtagtt  | ttatggtttt  | aggtcttata  | cttaagcctt  | taatccattt  | aaatttgatt  | 33540 |
| tttgtatgtg  | gtgagagaga  | ggggcctagt  | gttattttttc | tgcacatgga  | tatccagttt  | 33600 |
| tcccagcacc  | atttattgaa  | gaacctgtcc  | tttccccac   | tgaatattct  | tgaactccatt | 33660 |
| attgaaaacc  | agttggccgt  | gaatatgtgg  | atttatttct  | gagttcttta  | ttttgttcca  | 33720 |
| ctggctctgtg | tatctgtttt  | tatgctggta  | tcatgctgtt  | gtgggtacta  | tagctttgta  | 33780 |
| gcatattttg  | aagtaagggtg | atatgatgcc  | tccagttttg  | ttcttttttgc | tcagaaatgc  | 33840 |
| tttggttggc  | tgagtacagc  | agctcgtgcc  | tataatccca  | gcactttggg  | aggccgaggc  | 33900 |
| tggtagatca  | cctgaggtca  | ggagttcgag  | accagcctgg  | ccaacatggc  | aaaacctgt   | 33960 |
| ctctaataaa  | aatacaaaaa  | ttagccatgt  | gcagtgggtg  | gtgcctgtaa  | tcccagctac  | 34020 |
| ttgggaggct  | ggggcaggag  | aatctcttga  | acccgggagg  | aggaggctgc  | agtgcagcaa  | 34080 |
| gattacgcca  | ctacattcta  | ccctgggcaa  | acaaagcgag  | actctgtctc  | aaaaaaaaaa  | 34140 |
| agaaaaaaaa  | aaaaaagaaa  | tgtcttggct  | atttgaggtc  | ttctgtggtt  | ccatacaaat  | 34200 |
| tttaggattg  | ttttttctat  | ttctgtgaag  | aatgtcattg  | gtgtagagat  | tacattggat  | 34260 |
| ctgtaggtag  | cttttggtag  | tatggttttt  | ttcacaatat  | taattatttc  | agtcacacaa  | 34320 |
| tgtggtgtct  | ctttcaattt  | ttttgtgacc  | tcttcaattt  | cttacatcag  | tgttttatag  | 34380 |
| ttttccttgt  | aaagagggct  | ttcacctcct  | tgggttagatt | tattcctaag  | gttgtttttg  | 34440 |
| gaagttttta  | aaaaatctat  | cgagaatggg  | attgctttct  | tgatttcttc  | ttttgctagt  | 34500 |
| tcgttgctcc  | tatatagaaa  | tgattctgat  | ttttgtgtgt  | tgattttgta  | tcttgcaact  | 34560 |
| ttactgagtt  | tgtcagttct  | tagtttggtg  | gagcttttag  | ggttttctgt  | atataagatt  | 34620 |
| atatcaactg  | caaataggga  | cactttgact  | tcttcccttc  | agtttggtatg | ccctttattt  | 34680 |
| ctttcttttg  | cctagttgct  | ctggtcaa    | atgttgataa  | cttgttgatg  | ctgtcctcta  | 34740 |
| atgctcgcag  | cgtctgtgtt  | catgaaaccc  | tttgttgat   | tccaatgttg  | atagcattat  | 34800 |
| tcacagtaat  | caaaatgtgg  | aaacaaccca  | aatatctatc  | agtggatgaa  | tggataaaca  | 34860 |
| aaatgtggta  | tgtatgtgtg  | tgtgtgtgtg  | tgtgtgtgtg  | tgtgtgtgtg  | tgtgtgtata  | 34920 |
| ataaaatgtt  | attcagcatt  | aaaaaggaat  | gaaattctga  | tacatgcgac  | aacacaagtg  | 34980 |
| aaccttgaaa  | acatgctaag  | tgaaataagc  | tagttgcaag  | aggacaaata  | ttgaatgatt  | 35040 |
| ccacttaaca  | tgaaatatct  | agagtagtac  | acagattctt  | agagacagaa  | agtggttggt  | 35100 |
| aggttaccag  | cagctagggg  | gagcgggaaa  | tagggaatta  | ctgcttcatg  | gttattaaag  | 35160 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| agtttccctt  | tgggggtgatg | aaaaagcttt  | ggaagtaggt  | agtgatattg  | gttgacacaaa | 35220 |
| attgtgaatg  | taattatttg  | cactgatttg  | tacacttaaa  | aatgggttaa  | gtgattttat  | 35280 |
| gttatactt   | actacaataa  | aaaaagtctt  | taaaaatctc  | agcaagatgt  | ttttgtagat  | 35340 |
| atagacaagc  | ttattctgaa  | atttatatga  | aagggcaaaa  | gttctagaat  | agctataaca  | 35400 |
| atthagaaaa  | agaagaataa  | aatgggaaga  | atcaacctac  | ccaatgttaa  | ggcttactat  | 35460 |
| ttaggtagag  | taatacaagc  | attgtttcag  | gtagaggggt  | agacacacag  | atcaatggga  | 35520 |
| taggctgato  | tataggtaga  | tcccagaaat  | agaaccacat  | aaatatgtcc  | aactgattag  | 35580 |
| gatttgattt  | tgtttgtggt  | gtttttttga  | gacagggtct  | cactatgatg  | tccaggctgg  | 35640 |
| tcttgactcc  | tgggctcaag  | taacctctc   | acctcagcct  | cctgaataac  | tgggattaca  | 35700 |
| gggtgcacac  | gccacgctag  | ttccagctga  | attttttttt  | tttttttttt  | tttttgagac  | 35760 |
| aggggtctcac | tttgtcgccc  | aggccaaagt  | gtagtgggtg  | agtcacagca  | cactgcagcc  | 35820 |
| tcaacctccc  | tgggtcaagt  | gatcctcctg  | cctcagcctc  | ccaagttgct  | gggactacag  | 35880 |
| gtgcatacca  | ccatgcctgg  | ctaattcttt  | tttttttttt  | tttttttttt  | tttttggtag  | 35940 |
| agatgagatc  | tcctgtgtgt  | gcctagtctg  | gtctcaaact  | ctcaaactcc  | tgggctcaag  | 36000 |
| tgatcctccc  | gccttgggtc  | cccaaagtgc  | taagattaca  | gacatgagcc  | gttgcgccca  | 36060 |
| gccccagct   | gattttttgac | aaaggccctg  | caaaggcaat  | tcaatgaaga  | aaggatagcc  | 36120 |
| ttttaacgaa  | tagtgctaga  | gcagtgggac  | aaccataggc  | aactaaagga  | atctctaaat  | 36180 |
| ctcacatctt  | atataagaat  | taacgcataa  | ttgatcacag  | atctaaatgt  | aaaacataaa  | 36240 |
| actataaaaa  | ttttgaaaaa  | aaatatagga  | gaaaatcttt  | gaatcccaaa  | acgaaacaaa  | 36300 |
| agcttttttag | actcgatact  | aaaaacgtga  | ttcataaaag  | gaaaacttga  | taaatcgagc  | 36360 |
| cacataaaaac | caaaagtttt  | tgctctgtga  | gaaaccctgt  | tatgaagatg  | aaaaggcaag  | 36420 |
| ctatggactg  | ggagaaaata  | tttgcaaac   | atatttctgg  | gaaagcactt  | gaatctagaa  | 36480 |
| tataataact  | ctcaaaactc  | aacagtaaac  | aaacgaacaa  | ttcagttaga  | aatggggcaa  | 36540 |
| aagacatgaa  | tagacattta  | accaaagagg  | atatacatat  | ggcaaataat  | gacataaaaag | 36600 |
| atgatcggca  | tcaataacca  | ttagagaggc  | caggtgtggt  | tcacacctga  | aatcccagca  | 36660 |
| ctttgggagg  | ccaaggcagg  | cagatcacca  | cttgagggtca | ggagttcaag  | accagcctag  | 36720 |
| ccaacatggc  | aaaaccccg   | ctctactaaa  | aatacaaaaa  | ttagccaggc  | atgggtggc   | 36780 |
| atgcctatag  | tcccagctac  | tcaggatact  | gaggcaggag  | aatcgcttga  | acctgggagg  | 36840 |
| cagaggttac  | agtgagctga  | gattgtgcta  | ctgcactcca  | gcctgggcaa  | cggaatgaga  | 36900 |
| ctgcctctca  | aaataaaaa   | aaaataacca  | ttagagaaat  | gcaaattgaa  | agcacaatga  | 36960 |
| gataaacacta | cacacttatc  | aaaatgacta  | aaattaaaaa  | tagtaataac  | accaaattgt  | 37020 |
| ggagagagtg  | gggaaagatt  | ggatcactca  | tacattgttg  | gtgggaatat  | aagatggtac  | 37080 |
| agtcactctg  | gaaaatactt  | tgtttcttaa  | aaaactaatg  | ttatacttat  | catataaccc  | 37140 |
| agcaattgca  | ttcctaggca  | tttatcctag  | agaaatgaaa  | accagtgttc  | tcacaaatct  | 37200 |
| ctgtacaccc  | gtgttttatag | cacttgaaag  | caatatttaa  | cagctttatt  | tataatagcc  | 37260 |
| caaaactgga  | aacaagccaa  | atattcctca  | acaaatgaat  | gcttaaataa  | atgatgctac  | 37320 |
| atacacacca  | tggcctaata  | cccagcaaca  | aaaagaaatg  | aactgttgat  | atacagaac   | 37380 |
| acaacaactt  | tgataaatcc  | caagcaaat   | ttgctgattg  | aaaaagaaaa  | tcttagaaat  | 37440 |
| ttatatactg  | catgattgca  | tttatataac  | atttgaaaat  | gaaaatttta  | gatcttatat  | 37500 |
| acctatgcac  | atattgaata  | agtatcctga  | agcttcaaat  | ttctctgact  | tcttaggact  | 37560 |
| ctgatctcta  | gacctgtttt  | taaaggagct  | tatctgatta  | ggtcaggcct  | accacacag   | 37620 |
| gagtagggaa  | gatacagagc  | ggtgcccac   | aagaagggtg  | aatcttggg   | ggcattctta  | 37680 |
| gaattctccc  | tgctgtaggc  | tgcatttgga  | atagagcttg  | caggggtttc  | ggtagtttct  | 37740 |
| ccgatccgaa  | aaagggtgag  | ctctagggtc  | cagtcagtat  | aatttcagtt  | tggcaacatc  | 37800 |
| tcttgacact  | attaagatat  | agattatata  | tatttgttag  | gttattttca  | tgtaaacact  | 37860 |
| tgattagggt  | tctagttaaa  | aaaatttatt  | tatttactca  | ttttaagttg  | agattgtttt  | 37920 |
| tgtactgaca  | tacctctcca  | tattctcaat  | ctaattgaac  | atcgtgcttt  | gttgcttata  | 37980 |
| gctgacaaat  | gattaataact | ttttagattc  | ttagtgaaaa  | tcttgactta  | tgctgctgac  | 38040 |
| acttttccca  | gtctgtgcct  | ataagcaggc  | ttttgtgctg  | gagatggcca  | tacttggtgc  | 38100 |
| caatgccgag  | tttttttctc  | cccagttccc  | tattgtctca  | ccctctgaaa  | gtcactgaag  | 38160 |
| tggtttttaga | gctttcattc  | tagagggagc  | tcttccccac  | acttacttct  | aagaagtaat  | 38220 |
| aatttttaagt | ttcataatca  | aggagatgtt  | gattttcttc  | ctggttttct  | ggtgctcctt  | 38280 |
| atcttttctg  | gatcactaac  | ttaaatttaag | atccacttat  | gaatgctgca  | ttcaggagca  | 38340 |
| cttgtgaaaa  | taaaogatgg  | catgcccacc  | atthagtttg  | ccaagctttc  | cttgaaaacg  | 38400 |
| caggtaagtt  | ctattgcaaa  | cgtgtctttt  | tgctgttttg  | tacaaatgac  | tgactattc   | 38460 |
| actggtgata  | ggcccggtgc  | accgtgtgtt  | ctcagtcctc  | gcattgagagc | ctgggtgaca  | 38520 |
| ctgggttctct | ctgctttggg  | gaccgtatct  | ggctgtgcgc  | tatgtggagg  | gctgagctcg  | 38580 |
| tgtcttctctg | ggtaaataca  | tcattgttctc | aacaagggtc  | gcatgtttct  | ttattgctgt  | 38640 |

|             |             |            |             |             |             |       |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| tttttcccca  | gagacttagc  | ttttctaggt | gtgtagcatg  | gctcttaaat  | acacacat    | 38700 |
| agcaactgtg  | caaattcagt  | aaaaatggag | agacttattt  | aaggaagaca  | gtagggaaat  | 38760 |
| gtctaactag  | tgcagtgtag  | gccattcagt | taaatttcag  | tatggctttc  | gatgcaggca  | 38820 |
| aagccgtttt  | gccttttgatt | ctcatttttt | aaatttagatt | ttaaaaattg  | agattctcct  | 38880 |
| tttttaaatg  | agatttttaa  | aattgagatt | ctcctttttt  | aaatgagatt  | ttaaaaattg  | 38940 |
| agattctcct  | tttttaaatg  | agatttttaa | aattgagatt  | ctcctttttt  | aaatgagatt  | 39000 |
| ttaaaaattg  | agattctcct  | tttttaaatg | agatttttaa  | aattgagatt  | ctcatttttt  | 39060 |
| taaagcaaag  | taactgagga  | aatctaattc | tctacttgga  | aatttggtga  | agcattgggg  | 39120 |
| ctttaaaaat  | ggtaatacta  | aactcaagtt | ccttcactct  | gaaccaatcc  | caggattcct  | 39180 |
| ttccataatt  | aggaacagat  | caaaacctat | aactgctgag  | ggaatttgaa  | actgactata  | 39240 |
| aaaatctttg  | tagatgcagc  | taaaaaacia | acaaaaccca  | cagactgaaa  | agtcctgttt  | 39300 |
| ctgggagggg  | ggcgggtgtg  | ttggggccgg | ttgtgcctgt  | tttttgattg  | gttgttcctg  | 39360 |
| tggttgagca  | cccttccctc  | tgcccttgct | ttctgaaagc  | tatcctagtt  | gttctcccca  | 39420 |
| gtaagccggg  | gctcagccgg  | cttcagctat | gtccatagtt  | tgggggcgtg  | ggcggtaggt  | 39480 |
| attcaagata  | gaagatggtt  | ggacagaaag | tgtgtgaggt  | gggaagggtg  | agtggtaggg  | 39540 |
| ctctggctcg  | ttcccgactc  | tgcaacagag | gttgttactg  | taatgtggag  | ctaataataga | 39600 |
| agtctgcgtg  | ctgggtgttt  | catctactgt | tttatgtaag  | aggccagtgt  | tcacaagtaa  | 39660 |
| aggcaacaca  | gaattggtat  | tgagaactgg | gcattcattt  | agattttaga  | gttctgttta  | 39720 |
| tgggtgctaa  | caggatcgct  | atactgagca | agggatgtgt  | gaaacttttg  | ttcacttgga  | 39780 |
| tgtgcttttt  | ttcttataac  | agctttattg | agacatcatt  | tacataccat  | acagttcgct  | 39840 |
| catttaaagt  | acacaattca  | gtggtttttc | gtgtattcac  | agaagcgtgc  | agccaccctg  | 39900 |
| cactacagcc  | aattttagaa  | tattttcacc | acttcagaa   | gaagcccta   | gcagtcagtc  | 39960 |
| ctcattcccc  | cccttccgg   | caacccccag | tctgtatttc  | ctgtctctgt  | ggatttcgct  | 40020 |
| attctggaca  | ttttatataa  | attgagtcac | atgatatgct  | ctcacttagc  | atactgtttt  | 40080 |
| caaggttcac  | ccgtgtttag  | tatgtttcag | tacttattcc  | ttgtcatagc  | tgaataatct  | 40140 |
| tccttatatg  | tatataccac  | gttttggtta | tcattttttc  | tatcgatgga  | tgcttgggtg  | 40200 |
| gttttcactc  | tttggtcatt  | atgaatagtg | tcagatatac  | ttttaatggt  | tatatttatt  | 40260 |
| gtataataac  | agtgcataatc | cgtacctcct | ggtgttaatt  | gtctgtatct  | taagagtttt  | 40320 |
| aagaattata  | tggttgtgct  | gtaattctag | cacttaggga  | ggccgaggcg  | ggcagatcgc  | 40380 |
| tttagctcag  | gagtttgaga  | ccaggctggg | caacatagta  | aaacctgtc   | tctactaaag  | 40440 |
| tacaaaaaat  | tagctgggcg  | tggtagagtg | ttcctgtagt  | cccagctact  | tgcaaggctg  | 40500 |
| aggcacaaga  | atcacttgaa  | cccaggaggc | agatgttgca  | gcgagccaag  | attgcaccac  | 40560 |
| tgtactgcag  | cctgagcgac  | agagcaagac | tcctgtctcc  | aaaaaaaaaa  | aaaatttttc  | 40620 |
| tttatgatga  | aaccacacaag | gctgtacaag | aaccttaaat  | atagtttgaa  | acaaagaaaag | 40680 |
| acacatgctt  | ttaagtacag  | tttaattctt | ttcttttctc  | ttatttcatt  | tttttttttt  | 40740 |
| ttggagacag  | agtctcactc  | tgtcgcccag | gctggagtac  | agtggtgcaa  | tcttggctcg  | 40800 |
| ctgcaacatc  | tgccctcctgg | gttcaagtga | ttctcgtgcc  | tcagccttgc  | gagtagctgg  | 40860 |
| gactacagga  | acacttcacc  | acgcccagct | aattttttgt  | attttagtag  | agacagggtt  | 40920 |
| ttaccatggt  | gcccccatca  | ttctcagcaa | actatcgcaa  | ggacaaaaaa  | ctgaacaccg  | 40980 |
| cttgtttctc  | ttcataggtg  | caaactgaac | aatgaggacg  | catggacaca  | ggaaggggaa  | 41040 |
| catcacacac  | cagggcctgt  | tgtgggtggg | ggggaggggg  | aagggatagc  | attaggggat  | 41100 |
| atacctaattg | ttaaatgacg  | agttaatggg | tgacgacac   | caacatggca  | tatgtataca  | 41160 |
| tatgtaacaa  | acctgcatgt  | tgtgcacatg | taccttaaaa  | cttaaagtat  | aattaaaaaa  | 41220 |
| aaaaaaagaa  | aaatgtgtta  | ttcctttgtc | tcgaaagtcc  | cctgcttttc  | agcttttgct  | 41280 |
| ttgggtgtct  | cttgtcagct  | ctttaatgct | attgttccct  | tgggtttcat  | ctggactcct  | 41340 |
| tttccctggg  | caatctcctc  | cacaccatct | gggaacattg  | cctttggctg  | gacaggtaaa  | 41400 |
| tctttgtctc  | tagcagactt  | ctctcctgtg | atctccagag  | ccaggaaacc  | actagacttc  | 41460 |
| cttaaaatct  | attcagccag  | cagactttct | ggcccagagc  | agcagctgtg  | agtggacagc  | 41520 |
| aggcacctcg  | ttagaagctt  | ggtgttgag  | gtgtgcccac  | gcattcctgg  | gagccatgag  | 41580 |
| aggggagggg  | agtgcctgt   | gtcacctgag | tcctaagttaa | gagagatcac  | accagagctg  | 41640 |
| attcctaagg  | agcactgggg  | cgggggaagg | gggtgttcgt  | gggtgggaaa  | ggaaatgacc  | 41700 |
| agccctttca  | ggaactacca  | ggagttaggc | agagctgggtg | acagtgtctat | gagcaagaag  | 41760 |
| gtggagagct  | ggggagggcc  | attatgaaga | gactggagat  | tgcaggctgt  | gaacttgagt  | 41820 |
| cctgatgtga  | ctgggaattg  | ttgagatggt | ttaggccgaa  | aagtcaccga  | cgatcagtg   | 41880 |
| ttagaaagac  | ccactgggca  | gcagtgggga | gtgtggattg  | gaaggagcca  | ggcctaagtc  | 41940 |
| ctgggaggac  | ttaggaggtc  | caggagtcca | gggagtccag  | gtggaattct  | tgtgtagaaa  | 42000 |
| cccagggaag  | gggaccatct  | gatgacacac | atcccatagc  | aacttccttt  | cctcctctgc  | 42060 |
| acctccacag  | cagagtgtct  | gtttccacta | tacgaagatt  | tcttaatggc  | aaggatcgca  | 42120 |

|            |             |             |             |             |             |       |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttttatccc | tcagtgtcct  | cagtgtcatg  | tgctaatacag | ttggagacac  | tcctaaatg   | 42180 |
| tatgccagat | gagtatatag  | ttgttagtat  | aagggcttta  | cggtaggtat  | cagaactttt  | 42240 |
| tcttttttac | aacttttata  | actttttataa | aggagccgac  | acctttttct  | gatattgtca  | 42300 |
| gaacaattgc | tgaccagtgc  | tatgggctgg  | ctgaggttgg  | caaccaagaa  | tctgttatct  | 42360 |
| cagtctacac | agcgtggagt  | tagaagattg  | gattgttgtt  | ggttagaatt  | ttgtccagag  | 42420 |
| agttagaaga | tagaagaaaa  | gggagggctg  | aggacataaa  | catccaaatt  | ttacaaagaa  | 42480 |
| gaaagttcag | tgctcctgtc  | tatattgaaa  | tagttaagag  | ggaaatgatt  | gggtctgggg  | 42540 |
| aatgcttcaa | aataatatga  | tcatagggaa  | agtaagtgtg  | gtgtggatga  | aacaagggtg  | 42600 |
| gccttgagct | gataatttac  | taaagcagga  | tgatgagcgt  | gtgaggttca  | ttatagtgtc  | 42660 |
| ctttctactt | tttaaaaagc  | ttaagatttt  | ccaaaatgaa  | aaagaagctg  | agcatggtag  | 42720 |
| ctcacgtctg | taatcccagc  | actttgagag  | cctgagacgg  | aaggatcact  | tgaagccagg  | 42780 |
| agttcgagac | cagcctgggc  | aatatagtga  | gacccccatc  | cctactaaaa  | aagaaaaaaa  | 42840 |
| attacacaat | gaaaagaata  | atacaggtgg  | tagaagaaga  | aataaaggta  | tgagtatatg  | 42900 |
| atttggcctc | ctacagacag  | ctacaggagc  | atgcctctgc  | tgggcggggc  | tggagcatag  | 42960 |
| tgccacctgg | agctgaagtt  | ctgttcgcct  | ttacaggaag  | tacacagcat  | tgtacggggg  | 43020 |
| ttttaccagt | taacagtggg  | cagtctctag  | aacactcagg  | ggtcaggggc  | aaaactccct  | 43080 |
| ctccctcaaa | ccagcccacc  | tgtcttggtg  | agaatggaaa  | acggagcttc  | ccagcaagcc  | 43140 |
| tcctgccttc | accccagcgc  | acaggtgcat  | cagggctctga | gggctcctcc  | ctgtggggcc  | 43200 |
| tgggaagccg | cccttcagtg  | ggctatggct  | gcattgtcct  | ctgatcttgt  | agctgtgtgt  | 43260 |
| gggtaacagg | cacatcgtct  | ccccatatcc  | cagagagtcg  | tgtccacttc  | ctatctgttg  | 43320 |
| tggcatcttt | ccatcactgt  | caccctcagg  | ataatggcac  | aggaactaac  | agaggcacta  | 43380 |
| agagcagtga | tatatactgg  | agggaggggc  | ggggctgatg  | acaaaaatca  | ccttctgttt  | 43440 |
| cagtaagagg | ttggttaaag  | tgatgcctaa  | aatggataga  | tcaggaaata  | acaataaaaa  | 43500 |
| cattaatgat | aaaaaggcaa  | ctgccagaag  | aactgaaaag  | taacctagtt  | caagggtggg  | 43560 |
| gtccctggag | agaagccttc  | agaagagcag  | aggaagagca  | gggaacattg  | cttttcatga  | 43620 |
| ctagcccttc | agtgcctttt  | gatatttcag  | ctaagtatgt  | agttctctaa  | tgatttctaa  | 43680 |
| attgttaatg | agtgtgatag  | aataggaatt  | ctggagaatg  | gtcaaaaagga | attgctagga  | 43740 |
| gaaaattagg | gaaagaaaagc | acagagcagt  | attgaggtgg  | gaggttggga  | gcagacagat  | 43800 |
| ggtagagtgt | cgtttgcatg  | gcaggaatga  | gtggctgaat  | ggcattaact  | ggccatgggg  | 43860 |
| ctctgaccag | agcttgggta  | tacagtttct  | gtgccagcac  | aaaccagcat  | taggtgcgtg  | 43920 |
| gtgtgtgagg | cggggtgatg  | gtcatatgga  | tctgggtccag | agctgtgggtc | ttgacataga  | 43980 |
| gttccatgta | tgaactttgg  | tggtggtggt  | ggttttgttt  | cgttttttta  | gaaaaaaaaga | 44040 |
| agagttgttt | tatgtcacag  | ccaaactcaa  | ggaggtgaca  | gcacaagacc  | caggttagaa  | 44100 |
| tgtacttagc | cagggtctcac | ggaatcttcc  | atcctttgca  | aagccaagaa  | gagtagaggc  | 44160 |
| ctcaaggaga | atagaggagc  | acatgggctc  | cttctgggct  | gagctctcac  | ccttgggtggg | 44220 |
| gcegetggtc | ccaggggggtc | gttgggagcc  | tgccggagatt | ctgccagta   | catacatacc  | 44280 |
| ttgcgttttc | tccgagcctg  | acaggtgctg  | ccttgaagcg  | gaaagacttc  | acctcacagg  | 44340 |
| gtggccagct | gtcctttctc  | ttcctctccc  | tgctctgtgt  | gctcatgata  | atggatagat  | 44400 |
| agacatctcc | attggtagaa  | taatgcagcc  | cttctgcag   | tccatgggtc  | cacgcatat   | 44460 |
| ggtcagtga  | cttagatgag  | tgctacataa  | agtgttaatta | gttctgcacc  | acgggtgtct  | 44520 |
| accttctgaa | aaggaaactc  | tggtattgac  | ttagaaatgc  | cagttttatc  | cttccgacca  | 44580 |
| gaacgcaagt | cctgaatgag  | gccagaggcc  | aatgaaagg   | cctttcttct  | ctcttttcaa  | 44640 |
| atctgctgta | ttaaacagtg  | acttttccta  | gtcattcagt  | attcattcat  | tcatccattt  | 44700 |
| attcaagttt | ttgagtacct  | tctttgtgga  | tggcacatag  | tagatgcaaa  | aatacaatca  | 44760 |
| tttttaaaat | ggaaagcatt  | tgttatgtgg  | tagtgggaata | taagaaatgc  | catgcttcca  | 44820 |
| tatttttatg | gtaatcttgg  | actcaggttc  | taccacccat  | acttaattag  | aaattcattc  | 44880 |
| tcagcctctg | ttttggcaga  | gacccattct  | cttccctcag  | gaacatacag  | aaacctcagg  | 44940 |
| gctgaactga | gctcagagg   | tggggaagg   | ccctgggtct  | gatcttcagt  | cgtgtgctgc  | 45000 |
| tgcgccgat  | tccctggccc  | ctcactcttg  | agggcagcat  | cctcactctt  | gtgtggcaca  | 45060 |
| ggcttggaca | tagaatgaca  | cttaagcacc  | agagcctccc  | tagcctgcca  | caaggacacc  | 45120 |
| ctacattccc | aggcactttc  | agaatgcaaa  | acctccctcc  | cattccctcc  | ctttccctgc  | 45180 |
| ctccaggttt | gggcccctgc  | tctccttttg  | tttcaaagca  | ctgtcctctc  | ttcctctacc  | 45240 |
| cttatccttc | cccatctccc  | tcccactccc  | caaaagagcc  | aaactaatta  | tggggattta  | 45300 |
| accttccaat | taaaggcgag  | gatttttttt  | tttcaatgca  | gttcctgaaa  | tgctaaagaa  | 45360 |
| attgtccagg | gaaaggagca  | caggaggagt  | tttgttttgt  | tttgttttgt  | tttgttttgt  | 45420 |
| tttgttttgt | tttgttttga  | gacagtctcg  | ctctgtcacc  | caggctggag  | tgcaagtggg  | 45480 |
| cgatcttggc | ttaccgcaac  | ctctccacct  | cctgggtttg  | agcaattctt  | gtgcctcagc  | 45540 |
| cccccaagta | gctgggatca  | caggtgcaca  | ccaccacgcc  | agctaatttt  | tgtattttta  | 45600 |

|             |            |            |             |             |             |       |
|-------------|------------|------------|-------------|-------------|-------------|-------|
| gtagaaatgg  | ggttttgcc  | tgttgccag  | gctggtttcg  | acctcctggc  | ctcaagtgat  | 45660 |
| ccaccaccc   | cagcctccca | aagtactggg | attaggcgtg  | agccactgca  | tccagcctga  | 45720 |
| tgtttaata   | atttcttact | tcagagtttc | ctgaagtgtg  | gcctatggag  | actagcatca  | 45780 |
| gaatcactgg  | aggagcttaa | agatatagag | acagggatag  | ggcctgagat  | tctgactttg  | 45840 |
| ttgtaagcta  | tccagagttt | gagagccatg | gcctagttct  | tggaaatacaa | aagtgtatct  | 45900 |
| gctcttgtat  | gggaggtagg | aagttacctt | ttcattcatt  | cttctacaac  | agtaatttta  | 45960 |
| aagttaggaa  | ttggtatctt | catagacaga | agaatgcagg  | aatcctcttc  | tcaggatgtc  | 46020 |
| tgtctttaca  | cagccatggg | atatggtaga | ccaagtttgt  | ccaacccatg  | gcccgcgggc  | 46080 |
| tacatgcagc  | ccaggacggc | tttgaatgca | gcccacaca   | attcataaac  | tttcttaaaa  | 46140 |
| catgattttt  | ttattttatt | tttttttaag | ctcatcagca  | ctatcgttag  | tgttagtgtt  | 46200 |
| agtgtatttt  | atgtgtggcc | aaagacaatt | cttcagtggt  | ggcctaggga  | agccaaaaga  | 46260 |
| ttggacaccc  | ctatatagac | cataattatc | acacgccttg  | agagggagag  | agatcaatgg  | 46320 |
| tgagggagcg  | tagaggagag | cctgccagag | tttcagtggt  | ggccggggat  | gcttaaaaaa  | 46380 |
| ggaggcagtg  | atgtgctgag | ccagaaagga | cagggaggac  | ccccatgcac  | acaggcaggg  | 46440 |
| aggctctctg  | gtgacatctg | ccgtcaggta | cagtagtaca  | gtagtagagt  | gagcgggaag  | 46500 |
| acttttgggt  | gaagtgcac  | attagtgtat | gaggaagtgc  | aggctcgtta  | tatcttacct  | 46560 |
| caaaccctgg  | ggctcactga | gcttcagagt | tcagagtttt  | tcagaattga  | tgaacacatc  | 46620 |
| cagaaggccc  | atggggcaga | tacaccttag | acataattgat | gaaagtggag  | ttggaattgg  | 46680 |
| gttgtggaag  | tcgtaggacg | tgaattctag | gtttcttggg  | gtgtggagta  | tgggtgtctac | 46740 |
| tgtttggaaa  | cagagaaaga | tgacattatg | agttttattt  | ataaaataag  | aatctcttct  | 46800 |
| ttgtattaaa  | atagattttt | atataggaaa | ttttcattac  | cttcagtatc  | agaaaggccc  | 46860 |
| ccaaatctat  | ttcacctgtt | ccataataaa | ttgtgaaaga  | aagatttgat  | tgggatattg  | 46920 |
| tctttaaaaa  | tgaacccaaa | aagtatggga | aaatacttta  | attttagagt  | taacaacatt  | 46980 |
| tatagttaca  | atcatgaagc | aggtaggtg  | ctgcttcatt  | agccaagtgc  | ataaaagaaa  | 47040 |
| aaaaaatcag  | atcttactga | gttctgtgag | acagatcatt  | atttttctta  | ttaacatatt  | 47100 |
| cattttattca | gcaaagtgtt | attaagtgcc | aattaatgta  | tcagtggctg  | cttttaggcac | 47160 |
| tggagacata  | tggtcagtgg | gataagggga | agtctctgtc  | ttcaatcctc  | acagatctta  | 47220 |
| ttttccataa  | gtaagtgtac | aaatacgggc | tatgaattaa  | atagagcaag  | ataaagagat  | 47280 |
| gggggtgtgg  | cttcagttgt | gcacatgggt | cccagaaatc  | tgatagctcc  | atcatggagt  | 47340 |
| ccttcacttg  | agacttcaca | gcaggatgta | aaaagcaaca  | gccaggaaac  | tccttagttt  | 47400 |
| tcccaaagtc  | ctcccttaat | ctcgtttaga | gaattatgaa  | acacttccta  | tggcttcctc  | 47460 |
| ctttatccct  | tcttcagttg | ataactgaag | ggctaaatct  | gccaagtctt  | cctttgccaa  | 47520 |
| tgggtttgtg  | cgagcagttc | tccaacagcg | cttccactaa  | aatcttgtct  | tttacaagat  | 47580 |
| ggattttaca  | gttggttggc | attgtatgac | agtctgtgag  | tagggacagc  | ccaatgtatg  | 47640 |
| tttgtttaate | aagtttatgc | cttttacatt | gagtaatttt  | taattacctc  | gtgtaatttt  | 47700 |
| taacagcact  | ctccttcctc | ttcagaagtg | ttcagtttga  | acaataagtt  | tgtcatcttg  | 47760 |
| tcagtcaaag  | gttgaccaac | aaagactgac | cgcagtgggc  | ggggatttca  | gcagggttag  | 47820 |
| atgtttctaa  | gtggtcagtg | ctttcataaa | cattttcacg  | ttatgacctt  | tgcctcctca  | 47880 |
| cacaatcttg  | taactatgga | gacttagatg | ataaagattc  | aaactaagg   | atccctttta  | 47940 |
| agtcttgggt  | accagccata | ccaaaagctg | ttttatatga  | ggttctttat  | gattataatt  | 48000 |
| tggacatgaa  | catggagtgg | tcctgcttca | gcaagaaatc  | aggctaataa  | ccattagcat  | 48060 |
| ccagtgtggg  | taaaggcggg | ggctgccctg | aggagtgttg  | tgtggcacag  | ggcgcagtaa  | 48120 |
| caccagataa  | gtcctcagac | agtgccttca | gcattccagt  | gtttttggca  | tctagcaggg  | 48180 |
| agtaaaagta  | tctcatagat | tctttaagtg | taaacaccga  | ggacacactt  | aacgttttta  | 48240 |
| aagattgaaa  | tccagacagg | gctgggggtg | actctaactc  | gtttccttct  | ctctctttat  | 48300 |
| gtgttataac  | ttggattctt | cattccccaa | aatgaccttg  | tttatataat  | tctccagttc  | 48360 |
| tatctaaaaa  | ttttatgtga | ggaccatggc | tcataattccg | cagctttcct  | tctctcaagg  | 48420 |
| agctaggact  | ggttattgtg | ataaatgaag | gaaaagaaaa  | gaaagtgggt  | ttctgtttat  | 48480 |
| tgtatgaaat  | tgatgtaact | ttgtatttct | gattataaat  | ggtagccaga  | gggaggacac  | 48540 |
| aaaaacttca  | aattccaagc | cccataatat | attatggcct  | gaatagtgtt  | ggaattttat  | 48600 |
| tatttatgga  | atgtcaaaaa | gaaaataggg | ccatcaaaaa  | agcaaacagt  | ggaagtctat  | 48660 |
| tttctggccc  | ctgtgagttc | ggctctaaaa | acagtgcag   | gggttggggg  | ttggtgggac  | 48720 |
| taggaaagca  | gatgctcgtc | atttaccttc | ctgtgtttgt  | catgagaatt  | gtctgtactg  | 48780 |
| ggattgtcac  | cttgggtcag | aattgtttcg | tttatagtag  | cattgtctct  | caaagagggg  | 48840 |
| ctgagttccc  | taaaaaaaag | caggcagcgc | accacaacat  | ctggctagta  | gcacgtatat  | 48900 |
| tatcggtgat  | agtaaggaca | caattctcgt | atttaataca  | tttgagggtg  | tgcatttttt  | 48960 |
| cgtgggctaa  | tatccctgaa | atcatgatgc | atcttatggg  | cagtgaactta | aaaagcatta  | 49020 |
| tgtcatagtt  | aaattgacag | tgtttttctt | aaattataca  | cagaataatg  | ccttttttcag | 49080 |

21/154

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccaatgttat  | tttagagttg  | atgaaatatg  | ttgcattttt  | atattcttgg  | tatgaagact  | 49140 |
| attaactgta  | taataaaatt  | atttaagttc  | tgatgacgta  | gtgtcagatg  | atgttcagtg  | 49200 |
| tagcaaaacg  | cgaggcacaa  | ttttttgtaa  | aaaggacaat  | atcttgcaat  | cagtacacac  | 49260 |
| tgagtacata  | aatgttgacc  | aatgaagag   | cacagaatgg  | ggctgcccac  | ctggaggcag  | 49320 |
| agccatgaat  | tggtactgag  | gcaaacaagg  | agagctttgt  | gtgtgggtgg  | aggcgagggg  | 49380 |
| agggagcgcc  | aggcttcatg  | gaggcatgcg  | agggtctcca  | gggtcaaaca  | gacagaacct  | 49440 |
| cagaactggt  | gtgaagctta  | gaaggggctg  | caggaggctg  | tgctaaagcc  | tcagtgtatt  | 49500 |
| gctatgggac  | tggcaactaa  | acaagtaaac  | taaaggggtg  | aggtaacttt  | ccttaagtga  | 49560 |
| aaaaagtttc  | ccttaacttct | ttaaagtgtg  | agctctgttt  | tcagcttaac  | tcttgaaatg  | 49620 |
| agaagcgtga  | atccgtgggtg | gtttcatgac  | ctgcataatc  | acactaactg  | tttttctgtc  | 49680 |
| acagattgoc  | ttctctcagc  | acagcaattt  | tatggacctg  | gtacagttct  | tcgtgacttt  | 49740 |
| cttcaggtaa  | ttttgcttat  | acagactctc  | tctgttcttt  | tgatttaggc  | actagatcca  | 49800 |
| ttaatagagt  | atcagctcct  | aaaagataaa  | aagaataatc  | ctgtgtttcc  | accacagatt  | 49860 |
| aaatttactt  | catatgtgtt  | cctgaaaagt  | tagatgtcag  | ttacagtaat  | ataaactcaa  | 49920 |
| tctagcatga  | caaagtgaag  | acgtttgggt  | tgtatgttct  | ggcttccagc  | gggggggcttt | 49980 |
| gcatgtatcc  | tgtatttaatt | ttgcatgcat  | cctatagata  | atctggctct  | tcatatgggg  | 50040 |
| agaatttgga  | ctaaaacaac  | atttaattaa  | taatgggttac | tgatgtgagc  | tacacatatg  | 50100 |
| cttatgcttg  | ttcaagagat  | gattctgggtg | attatccaaa  | gtttttaaagc | tgaaagtgat  | 50160 |
| tttttagtgc  | cctgggatttt | tcttttattt  | ccatagtttt  | tttcgtagtc  | taattttatc  | 50220 |
| tttatgaaaa  | tactatagac  | aaataattga  | aaaagacaaa  | aaacgtgctg  | cagtttttaa  | 50280 |
| aacagaaagc  | agcaatagcc  | tgccctgcgc  | cttccccacc  | ccacgtcccc  | caattttaat  | 50340 |
| tgttttgttt  | tttacttctc  | tgtgtttaca  | cttgggtttt  | cagttaagtt  | ctcttttgat  | 50400 |
| gtcctactgt  | agaagatgag  | gatatactgt  | gttacatggc  | ctccccctagc | atctccccctc | 50460 |
| acaccacaaa  | agcatacaaa  | agccccatct  | ctttcattta  | ataccattgt  | atcataatgc  | 50520 |
| atggctaaat  | cagtgatccc  | tatctatgtt  | atggccataa  | aactcttatt  | cacagctgca  | 50580 |
| ttgaacagtg  | attacactac  | cacatttttt  | gggtagaagg  | atcattcttt  | gcttatccta  | 50640 |
| aatgtagtat  | ttgccatggt  | ttggcttttt  | aattgtttga  | tttttctatg  | tgccctacccc | 50700 |
| tgatttatca  | ctatactctt  | gctgatgtac  | ctttagaaca  | gtaaaaaata  | aaataaaaata | 50760 |
| aagtattgac  | gtactgtgac  | agaactgaga  | aatctcacia  | tggtcacaca  | caggaaggtg  | 50820 |
| ctctcttggg  | ttgaatttct  | ctcgagttg   | cccacctag   | aaccttccca  | cctcttgccc  | 50880 |
| ccaagtgcac  | ttgttctctga | ggcctgctgc  | acaactgtca  | tttggaaact  | tcctttgccc  | 50940 |
| tctctctata  | ttgaagtctc  | tgtttcccag  | atccccatgtc | ttcaccttcc  | tatccattcc  | 51000 |
| ctctttttac  | tggaagcacat | tctctagatt  | tatgagaaaa  | gatgcctggg  | agataacttt  | 51060 |
| tctgtgagct  | ctcctgactg  | aaaatatatt  | tattttatcc  | tacaatttga  | ttgaatttct  | 51120 |
| gggggcaggt  | gctatggctc  | acacctgtat  | ttctagcact  | ttgggagacc  | aaagtgggtg  | 51180 |
| ggctcacttga | gctcaggagt  | tcgagaccag  | cctggacaa   | atggcaaaac  | cttgtctcta  | 51240 |
| caaaaaatac  | aaaaattagc  | tggttgaggt  | ggcgcaagcc  | tgtagtccca  | gctaatacagg | 51300 |
| aggctgaggg  | aggagaatgg  | cttgagccta  | ggaggcagag  | gttgcaagtga | gcggagattg  | 51360 |
| caccactgca  | ctccactctg  | ggtgacagag  | tgagacctctg | tctcaaaaaa  | aaaaaaaata  | 51420 |
| aaagtctgac  | tacagaattc  | taggttcaaa  | atcattttcc  | attagaattt  | agaagacgtt  | 51480 |
| atactatttt  | ttctagtgtc  | cagtataata  | ttgagaaggg  | tagtgctgtt  | ataattccca  | 51540 |
| ctctttgtat  | gggatccgtt  | ttctctctta  | aaaagaaaag  | tataatttaag | tgtaatatata | 51600 |
| agttacagaa  | acatttgcaa  | atcctatgtg  | tgtagctcag  | tgaattatca  | caaattgaac  | 51660 |
| acttctagag  | tattgataat  | attatctctt  | gttcacccag  | aacacttcag  | gcttcttatt  | 51720 |
| tcttctctgt  | cagtttccgg  | aagttgtata  | tagttttctt  | ctaattgatcc | cattccatct  | 51780 |
| gttccattct  | cttgttcact  | gagaacaaga  | gttaatatata | ccaatactct  | agaagttcct  | 51840 |
| cttggtgttc  | ccctcactca  | caacctcctg  | cctcttcccc  | aaaggtacca  | tttacctgac  | 51900 |
| ttctcaaaac  | atacattgca  | tttgctattt  | ttgaaatgca  | tatgaattat  | ttcttttttt  | 51960 |
| tttttttttt  | ttgaaactta  | gtctctctct  | gtcaccaggg  | ccagagtcca  | gtgttggtgat | 52020 |
| cttggctcac  | tgcaagcaac  | cgccgcctcc  | cggtttcaag  | caatttctct  | gcctcagcct  | 52080 |
| cccaagtagc  | tggtgattaca | ggcacgtgcc  | accacgcag   | gctaattttt  | gtatttttag  | 52140 |
| cagagacggg  | gtttcaccag  | gttggccagg  | ctggtctcga  | actcctgacc  | tcaagtgate  | 52200 |
| caccgcctc   | tgctcccaa   | agtgtctggg  | ttacaagcat  | gaaccacccg  | gctcagtcac  | 52260 |
| aatgtatata  | aatgaaatca  | caacataaat  | tttaagatga  | tttgggtgat  | gaagaaaata  | 52320 |
| ctgaaaactt  | gtgggttgca  | gttaaagttg  | tgcttcgagg  | ttatagctgt  | aaatctgtgg  | 52380 |
| ctctccacct  | tagctgcacc  | tagaatcact  | tgccgagctt  | ttgatgtcca  | agccactctt  | 52440 |
| cagaccattt  | gactcagaac  | ctctaggggt  | gcaccccaaa  | catttgaaaa  | ctctccaggt  | 52500 |
| gattccaatg  | tgagctaaag  | gttgagaatc  | actgctttaa  | aagcacatat  | tagaaaagaa  | 52560 |

|             |             |            |             |             |             |       |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| aaaacattcc  | tctcaataac  | tcagaactgc | taattaaacc  | taaagaagta  | gaagaaagaa  | 52620 |
| aaatataaat  | aagagcaaaa  | attaatgac  | taagggaaga  | agtacagtag  | agaggatcct  | 52680 |
| aaaaactgtc  | agttggatct  | ttgaaaagag | taagaataat  | attgacaaac  | ccctggcaag  | 52740 |
| attaaccaag  | gaaaaaagag  | agaaggtata | tggtatatgt  | agccaatgtc  | agaaataaaa  | 52800 |
| aagggggcat  | tcctccaaat  | cccgcagacg | tttttttccct | ccagaagggt  | atttgatgaa  | 52860 |
| caactttatg  | tgaatgaatt  | tgaaaatgaa | acagatggaa  | tgggaccatt  | aaaagaaaaa  | 52920 |
| tatgcaactt  | gccaaaactg  | acacagaaga | agtaaggagc  | ctgaatagtt  | ctgtatccat  | 52980 |
| ttttaaaaaca | acactaattt  | aagttttgtc | atagtgaaaa  | ctctaggccc  | caggtggctt  | 53040 |
| cattggtaac  | ataacctttt  | taggcggggt | gcagtggctc  | acgcctgtaa  | tcaccagcact | 53100 |
| ttgggagggc  | gaggtgggtg  | gatcacgagg | tcaggagatc  | gagaccatcc  | tggctaacac  | 53160 |
| ggtgaaaccc  | catctctact  | aaaaatacaa | aaaattagcc  | aggcgcgggtg | gtgggcgcct  | 53220 |
| ataatcccag  | ctactcgggg  | ggctgagaca | ggagaatggc  | gtgaacctgg  | gaggcagagc  | 53280 |
| ttgctgtgag  | ccgagatcgc  | gccactgcac | ttccacctgg  | gcaacagtgc  | gagactccat  | 53340 |
| ctcaaaaaac  | aaaaaggaga  | atactgtgat | catcttaatc  | aatggaggag  | ataaagcatt  | 53400 |
| tgataaaaatt | taatacccac  | tcatttttaa | aaaaatttct  | tagaaaacta  | gaaagagaaa  | 53460 |
| cttctggaac  | actgataatg  | gaacaacata | aaaagtcaac  | agcaaaccac  | tttcttaatg  | 53520 |
| gtagaatgtt  | agaagtttga  | gtttgggaac | aagacagaga  | tccactttcc  | acagtttcaa  | 53580 |
| ttactcaggg  | tcaaccgtgg  | tccaaaaaca | ttaagtggaa  | aattccagaa  | ataaacaatt  | 53640 |
| tataagtctt  | aaattgtact  | ccattctgag | tagtgtgatg  | aaatgtcgtg  | ccatcctgct  | 53700 |
| tcgtcccacc  | tgggatagga  | atcatccctt | ggtccagtgt  | atccacgccc  | ctatgctacc  | 53760 |
| caccactag   | tgtcttgggt  | atcagatcta | aaaatcatag  | tatatgtagg  | gttcagtact  | 53820 |
| atccatgggt  | tcacgcaccc  | actgggaggg | tggaatgta   | acccacaga   | ttaagcagga  | 53880 |
| cactatattg  | ccaaatccaa  | aggatacttt | ttggttctcc  | tcttcttgac  | ttctcggtga  | 53940 |
| tatcaaccaa  | ctgactgctt  | catgagcgag | actcttccag  | ccgtctctga  | cccatgctgc  | 54000 |
| cctactgctg  | cttcttagcc  | tgctattat  | gcaatctttt  | ctattcagcc  | tctcactggt  | 54060 |
| gattttttcc  | cagaccttga  | ccctggccgt | cttctcactc  | tgtgttctct  | gactggactc  | 54120 |
| tgcttcccc   | atctgcagca  | cgacatgccc | agaagtgacc  | aattccatcc  | ttcctcaccg  | 54180 |
| gtgttcccc   | cctcactaac  | tggcaacatg | gctctgctgg  | ttccaagcca  | cttctcttgc  | 54240 |
| tttctactcag | caggcactcc  | cttaccaggg | cccagagctg  | tggccgtcag  | tgcactctcc  | 54300 |
| acctccactc  | ccacccccac  | caaccctagc | atcacagtct  | cgccctggacc | ttttcacttg  | 54360 |
| actttggccc  | ctccagaatc  | cacaaaaagt | ccagagaaga  | cctttaaaga  | tcatgtatct  | 54420 |
| agccccttac  | cttctgttgg  | tatctaaaat | aaaattcagc  | gtcctgggtg  | aactggcgct  | 54480 |
| cgcttttcca  | tattcagctt  | gcactcttct | cttccacatt  | tctgctacca  | ggaccatctc  | 54540 |
| ctatgtcttc  | aaaagcacaa  | gctctttccc | tacactgagc  | atttgtacat  | gcttatccca  | 54600 |
| ttttctggaa  | cactcttctc  | actacccttc | attctgtata  | ctttcatacc  | ttgcttgatt  | 54660 |
| tttccagaca  | acaacaaaatt | ctccaattct | ctgacaccaa  | ctcagtgtct  | cacagttcag  | 54720 |
| ttccattcag  | acactaccca  | tagttggagc | aaatcctgca  | ggttaggggc  | tcagtcccac  | 54780 |
| aagactgctc  | ccacttcaga  | cactaagcac | aatgggtcc   | ccaggctacc  | tcactgctac  | 54840 |
| tgggccaact  | acaaaatcag  | ggcttctggt | gactgctcac  | tccaagattt  | gacaatttgt  | 54900 |
| tagaacagct  | cacagaactc  | aggagagcac | tatacgtaga  | attacagttt  | tattatacag  | 54960 |
| ggtacaactc  | aggaaaccgc  | aaatggaaga | catgaatagg  | acaaggcatg  | gggggtgggtg | 55020 |
| agcacagctc  | tttcatgccc  | tctctggaca | agcctcctcc  | ctgcacgtgg  | atatattcac  | 55080 |
| tatcctggaa  | gctcccccaa  | gcctcatcat | tctgagtttt  | tttattgagg  | tttcattaca  | 55140 |
| taggcagtag  | ggatgaaatc  | actttctatt | ggtgactgaa  | ctcatctcca  | gcccctttcc  | 55200 |
| ccttcttggg  | ggttggggga  | tggagctgaa | agttctaacc  | ctctaaccac  | ctagttgggt  | 55260 |
| tttctggtga  | ccagcccccg  | catctgaggg | tcttttagggc | cctctcccca  | tgagtcattc  | 55320 |
| attagcacac  | aaaagacact  | cctgtccctc | tgaaaattcc  | aaggggtaag  | ctaaattgtc  | 55380 |
| tccccttgtt  | cctcactcct  | gtagaactgt | attatttttc  | atcgtggtat  | tcatacacaat | 55440 |
| ttgaagaata  | tattcatatg  | tttacttgct | taatatctgg  | gtcggtagcc  | caagcttcag  | 55500 |
| gggagcagtg  | cactgtctct  | cgttcactcc | tatacacaca  | cagcttcagc  | atctagtaca  | 55560 |
| taatctagta  | agatatttgt  | gggaagtaaa | cttagactct  | tcagggttaag | tgcagagtga  | 55620 |
| ggaggaacca  | taggcaccaa  | cctctcaggg | aaaaaagctt  | gtgtcaatgg  | cttgctttta  | 55680 |
| agggatgtag  | gggctgtatt  | gtagctccag | tagacatcct  | taaattcttg  | gagcatgtgt  | 55740 |
| atggaataag  | atggccctgc  | tgccaagtac | gtacagcctg  | gtgaggaccg  | gccacttcta  | 55800 |
| ctggacattc  | agtccctggg  | cgaccgaatg | acactatcag  | aataagtcca  | ggaagtactt  | 55860 |
| ttcaggccca  | tggcctcttg  | tgtagtaaac | aaggacagat  | ggcattttca  | tcacttctcc  | 55920 |
| cttgtggggg  | ttgaacttct  | ggaaaagaat | ccggaacttt  | ggtttttcta  | gacgatgttt  | 55980 |
| actaactage  | gatgataaac  | ttttagctaa | gaatagagtt  | actttcatgg  | catggggaag  | 56040 |



|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| aaaatatttt  | cccaggacaa  | gttaaccatg  | tctaaaagaa  | tttattttagc | tttttttttt  | 56100 |
| tttttttttt  | tttttttggg  | ataggggtctc | tgttgcccag  | gctggaacac  | agtgggtgcag | 56160 |
| tcatagccca  | ctgcagcctc  | aaactcctga  | gctgggtctca | agcagtcctc  | ctcacctcag  | 56220 |
| tctcccaagt  | agagctggga  | ctgcaggcat  | gcaccatcat  | gccagctaa   | tattttcttt  | 56280 |
| tttgtagagt  | tgggttctca  | ctatatgtcc  | caggctggtc  | ttaaacttct  | ggcctcaagc  | 56340 |
| agtcctccca  | tctcagcctc  | ccaagggtgct | gggattacca  | gcgtgagcca  | ctgctcccag  | 56400 |
| cctgggtccaa | aagaatttaa  | aatggagcct  | ataaggaaaa  | ggggaaagg   | atccagataa  | 56460 |
| tcctaagcac  | tctgaagaat  | ggctcatata  | atatatggaa  | aaaagaacaa  | aagtcggggg  | 56520 |
| gaaattcgat  | gattcaaatg  | agaaaatcaa  | tgtaaagtga  | gttagaaaaa  | atatttccag  | 56580 |
| actccgggca  | cagaaaagta  | aactctagg   | cctgacacag  | tctgttgaga  | ctgtgtccta  | 56640 |
| atgagatgct  | tttccagcct  | caggaaggag  | cctccttaaa  | gagggaaata  | agagactctc  | 56700 |
| ctctgcttgc  | ttttctgtag  | cttcttgtag  | ctggagcagc  | aagtagaagg  | gactggggaa  | 56760 |
| ctggtgttgt  | caacagatct  | tcaactgaata | cagaagggcc  | ctctcacctt  | ccctgaggca  | 56820 |
| tagagggaag  | caggcacaga  | ctgtattacg  | ggtggaccca  | agagagggtt  | ctgagtgaca  | 56880 |
| ccagggtcct  | cgctctcag   | cagcagcggc  | aggctcttga  | ccaacttaga  | agctgttgcc  | 56940 |
| gttggcttca  | tgctcttctc  | tgtggatttg  | ctgcctcctt  | atgtttatct  | attagaataa  | 57000 |
| aaataacaaa  | aatcaggatt  | atgctgtgaa  | gccaaagaa   | actgagaagg  | ttattctcag  | 57060 |
| acctctttac  | tggttcttcc  | tgttctctc   | aacttctcat  | tgttaggggtg | gccccaggac  | 57120 |
| attctcatct  | atctgtattc  | actctcttgg  | tgaatgtggc  | cagggtttatt | gcttgaaata  | 57180 |
| ccatatatat  | ccttatgcct  | tccaaatgcc  | atgcccagcc  | cagacctggc  | agccagcctc  | 57240 |
| caggttagaa  | tatctgccc   | ggacatctcc  | cctgggtgtg  | ccggtagata  | tctcagggtt  | 57300 |
| ctgcatgtct  | aaaatggatc  | tactgatatg  | tccagattcc  | aaagtgtttt  | tcacacctct  | 57360 |
| gcttccacag  | ctctggtttg  | agccaccatc  | atctcaccta  | gactgctgg   | ctccctgata  | 57420 |
| atgccattgc  | acctctaattg | gttttctcag  | caaacgttaa  | atcacacac   | tctgtttaaa  | 57480 |
| aaactgcact  | ggctctttcc  | caatgccctc  | ataatgcttg  | cagtgcacca  | catgatctgc  | 57540 |
| ccttcacctc  | caccttctc   | tctgcctttg  | tctctggcct  | tacacacccc  | actccaaaca  | 57600 |
| tgctggcctc  | ctgaggctcc  | ccaagccatc  | ccaggcacgg  | ccttgccctg  | ggctttgcac  | 57660 |
| tgactgctcc  | tttttccaga  | cgcacttttg  | caggcgtcca  | tgagccacct  | cccatacctt  | 57720 |
| ctccaagtct  | tgttccaat   | gtcaccttct  | gagaaagaaa  | ggccacaca   | ctctgacct   | 57780 |
| cctgctgtgg  | agtgcactg   | gtactcctga  | accccttgct  | cttagcatct  | ttttatttct  | 57840 |
| gtagacctta  | cctacctttt  | ggcatgctct  | gtaatttgct  | tatttactta  | ctgtttgtct  | 57900 |
| ccacctgctg  | gaatgtaaat  | tccatgggac  | agtaatcttt  | gctctgttta  | tcagtgtttt  | 57960 |
| ccatgagcct  | agaacaggac  | ctgacccaat  | taaagtattg  | attaaaaaga  | caatgtaggc  | 58020 |
| cagtcacgtg  | aggcctgtaa  | tctttgccaa  | ggcaggcaga  | tcacttgagg  | ttaggagtgt  | 58080 |
| gagaccagac  | tggccaacgt  | ggtgaaacct  | cgtctccact  | aaaaatacaa  | aaattagctg  | 58140 |
| ggtgtgatgg  | cgcacacctg  | tagtccaagc  | tactcaggag  | gctaaaagag  | gagaatagaa  | 58200 |
| tcgcttgaac  | ctgggaggca  | gaggttgacg  | tgagctgaga  | tcgtgccact  | gcactccagc  | 58260 |
| ctgggtgaca  | gagcgagact  | gtgtctcaaa  | aaaaagacaa  | tgtagactgt  | cagaagagtt  | 58320 |
| ggcagtgctc  | tgatgcaga   | gaggcccaga  | gatgaaggct  | ttccaccaag  | tgacgatcct  | 58380 |
| gggggggttg  | gaaggcaatc  | ctattgaaat  | gcctgtgtgc  | agggcaccca  | tgggaggagt  | 58440 |
| ctgcaccctt  | gtacccctta  | gggaggaggc  | agtgagcctt  | gagataagg   | cagttcagtc  | 58500 |
| atcctagcat  | cttagagtcc  | ctctctccag  | ccagcccagg  | agggctgtgt  | tgggtgcagc  | 58560 |
| agttgggtgc  | actgaggtgc  | cattcatgaa  | gctgacccca  | tcctacaaac  | gcagagctca  | 58620 |
| gccagcaaga  | gagtaccttg  | atgacactgg  | cagagagcct  | tatgtctaca  | tgtagcactc  | 58680 |
| tggaaaaata  | ccctgtagct  | agcccctttt  | acaagttaat  | gagtaaaaaat | gcagttatag  | 58740 |
| cagttatcta  | tttttttact  | tgttaagctt  | attcttctct  | ttaaccgcaa  | gtcagctggt  | 58800 |
| ctctttgcca  | tcactgtagc  | ttttcaaate  | attccacacc  | tgggtttatac | atgagagagc  | 58860 |
| attgaaatcc  | tctgcaacta  | ttaaggccac  | tgggagttat  | tttctcttaa  | tatttttaag  | 58920 |
| aaagttgttg  | gcttgggttt  | gatcatttca  | aatgtgtttt  | attttactta  | aaaatatgta  | 58980 |
| tttgtaaaaa  | tattgaatat  | acactaataa  | ttacacacat  | ttcttctgtt  | tgcctcttcc  | 59040 |
| ttatgaaaag  | ttgattgtta  | ctttaataat  | taactttttc  | accttcttag  | attttaccag  | 59100 |
| taaatatctg  | ctaacttgaa  | taaacagcat  | gactaatatt  | tgttataatt  | atcttagaat  | 59160 |
| tatttttctt  | atgtgggata  | agccactgtg  | ggccattctg  | atgaaggcca  | agtggaagg   | 59220 |
| cagggtctct  | tctgaatacc  | acatgggaac  | gatacagctc  | agaggcctgg  | ttagggaagg  | 59280 |
| tggcttacgg  | aagttgtctc  | actcattacc  | tcatgagccg  | gcacattgtt  | gtgacaaagc  | 59340 |
| atcagaaggg  | agcccggagg  | tggcatgtgg  | gagctgtttt  | cactgcctgg  | gcctgagggtg | 59400 |
| tgggcaacgg  | tgggagctga  | ttctgggggc  | tgggataata  | actcagttct  | ttcctttctc  | 59460 |
| acagttgttt  | cctctctttg  | ctcctggtgg  | ctgctgtggt  | ttggaagatc  | aaacaaagtt  | 59520 |

|             |             |             |            |             |            |       |
|-------------|-------------|-------------|------------|-------------|------------|-------|
| gttgggcctc  | cagacgtaga  | gaggtaagct  | tcagtgggta | aagattaaag  | aatccctgga | 59580 |
| agagcttttt  | ttccttcttt  | ttctcttaag  | caagtgggtt | ttagctattt  | agtataatg  | 59640 |
| gacagacaga  | catctccacg  | gaaggaataa  | tgcagccctt | tctgcagtcc  | atgggtctgc | 59700 |
| agcatatggt  | cagtagactt  | agatgagtgc  | tacataaagt | gtaattagtt  | ctgcaccaca | 59760 |
| gtgtcttacc  | ttccgaaaag  | gaaactctgg  | tattgacttg | gaaatgccag  | ttttatcctt | 59820 |
| cagaccagaa  | cccaagtcct  | gactgaggcc  | agagatccaa | gggcgtatta  | gcacagccac | 59880 |
| cgcgtgatta  | ggcaccgcct  | tcatgagaac  | ttgtcatgcg | agagggaaat  | cagccattta | 59940 |
| agcatctcaa  | aaatttttca  | ttattcaagg  | aaagataaat | gtgtgtgtag  | ttaagttttt | 60000 |
| aacttgagtt  | ttatttttta  | ataagtcttg  | atttgtttcc | agagcttcat  | tccaaaagta | 60060 |
| ataagcaaat  | atttacagct  | gacaaattta  | aataaataaa | cttggttgca  | aattacaaca | 60120 |
| tatcaaatgc  | ccatcaatca  | atgagtggat  | aaagaaaatg | tgggggtgtg  | atgtgcgtgt | 60180 |
| gtatgtatat  | atataatatt  | atatacattt  | gcagcaacct | ggatggaatt  | ggagaccgtt | 60240 |
| attctaagt   | aaataattca  | ggaatggaaa  | accacacttc | atataattct  | acttacaagt | 60300 |
| gggaactaag  | ccctgaggat  | gcaaaggcat  | aagaatgata | cagtggattt  | tgaggactcg | 60360 |
| ggggaaagag  | tgggaggggg  | atgagggata  | aaatactata | catcgggtac  | agtgtacact | 60420 |
| gctcgggtga  | taggtgcacc  | agaatctcag  | aaatcaccat | gaaagaactt  | attcatgtaa | 60480 |
| ccaaacacca  | cctgttcccc  | caaaaacctt  | ttgaaataat | aaataaataa  | ataaataaat | 60540 |
| aaataaataa  | atatttttaa  | aataaccacc  | attaaaataa | ttacaacata  | aaaactttgt | 60600 |
| gtattgcaac  | tttcaattgc  | atcattactt  | tgcctttcat | ttagtcatgc  | ctgtggcatt | 60660 |
| gttggcattg  | ttgtcaatgg  | aaatgctcac  | gtgtatataa | gatctagaca  | aagaagctgt | 60720 |
| ctgggggttt  | tttggatcta  | ctcatgcctt  | tttctatata | aaaatgtatt  | tatagataag | 60780 |
| atttagaaca  | gaagggaagt  | atgtaaaatc  | acaactcaca | agcactcagc  | actgttgata | 60840 |
| ggatattcat  | gtctgaatag  | ggttaagaca  | ggctccagga | tatgggcctc  | ttgttgtggg | 60900 |
| ccacagtacc  | accttttctg  | acccacataa  | agatgatgtt | atcatagagg  | gaatgttgca | 60960 |
| cactctttat  | tattattttt  | taaataagt   | gataacatat | tcaaaggggt  | ttttagataa | 61020 |
| ctgtattact  | gatatacttt  | aaaagcctaa  | tgaattacgt | gatgttctga  | aaaaaattat | 61080 |
| attttcaatt  | taattttaat  | gtatatttat  | tcacaatgga | gtaagaaaga  | agtaagaagc | 61140 |
| cacagtaaac  | cagttatcaa  | gaaaggagtt  | gaaaactttt | gagtgtact   | ctctccaaga | 61200 |
| aggacattag  | accgagggag  | tttcaaggtt  | gagatgagct | tatggagatt  | atgatagcct | 61260 |
| gttctctcct  | gagtgcctat  | gagaaataat  | gttttctaaa | atctgtaaag  | caaaaggaca | 61320 |
| gatggaaagc  | ttgaaagagt  | ccgaattgat  | tttatacaca | taaatcttgg  | taccaaacc  | 61380 |
| tgacaaaggt  | agcaccaaaa  | tacaaagggc  | agctgcatag | tgaatgtatg  | tgcagcagtc | 61440 |
| taattgcagt  | gtaatgaaag  | agtcacttaa  | taccagaaaa | cctattagaa  | tatatcatgt | 61500 |
| caaaaaaaga  | atataatcat  | tcaatgggga  | aaaaataaca | cctgaattct  | tcctagtggc | 61560 |
| ccagaaatca  | tttgcgtcatt | tattgaaact  | caatatccat | tcctaatttg  | aaaacaaaac | 61620 |
| tcacagtaaa  | atcaaaaata  | ccttttttaac | gctatgaaga | atatctggat  | aaccagcatc | 61680 |
| actcatgata  | aaacaaacaa  | acaccagagg  | ttgtatccag | gctagttagt  | tttaacaaga | 61740 |
| gctgacatct  | gagcatttaa  | aaattccagg  | cactgagttc | taagtgggtt  | tacgtagatt | 61800 |
| aactaatcct  | tatctcaaac  | ctatgaaata  | ggttatgtcc | ctcgttatgt  | tttatggatg | 61860 |
| aggggactga  | ggctcagaaa  | ggtgaagtga  | cttacaccac | atagttaact  | aagaggcaga | 61920 |
| cgaggcttta  | actagggact  | tctgtcttca  | cagcccaagt | tcgttatctt  | cgtgtaatcc | 61980 |
| tacagtcat   | cagctttgtg  | ctagatgttc  | tagactgtgc | actaagtga   | agggcactaa | 62040 |
| gaagtgttgt  | tactattgga  | aaggaaacaa  | aattattatt | ttcatataaa  | gttattaaaa | 62100 |
| caatgtagta  | gagatttgcc  | tgcataaaaa  | tgaaggctta | agtatatcaa  | aaaataaatg | 62160 |
| aataacacag  | aagctaaaca  | atcaactgga  | aattgtttgc | agtgtgcatg  | agagaaagta | 62220 |
| agatgcacag  | cctacaggaa  | aaagattaaa  | atatatatat | gtgttttagta | tggagttggg | 62280 |
| catctaataa  | actatgcttt  | taccctgtgt  | atttggaatt | tctctgtaac  | tgtgaagctg | 62340 |
| accattcaga  | atataattca  | tcattttttaa | aatgtcagcc | attattttatt | cctctcaga  | 62400 |
| attgtagaag  | caatgatttg  | atgttatttt  | taaaaacaat | agtagtagcg  | gctgacgctg | 62460 |
| tttagtcctt  | gctacacacc  | agccccggtt  | ttaaagtgtt | ttacacatgt  | tgccctctga | 62520 |
| accctgtctt  | attgtgaccc  | cagcacttaa  | gcacagaccc | caaggagcat  | tgctgcatat | 62580 |
| attagtcogag | taaattggacc | tacagcttta  | aaaaatgtaa | gaatgttggg  | attttaagaa | 62640 |
| aatattttata | aaaatgcctt  | tctagaacag  | agccatatct | atccaggcca  | agattctgtc | 62700 |
| ttctacagta  | gttccaagga  | ggcagtttga  | agtaaggcct | acttgatatcc | caaggcataa | 62760 |
| cctcaaagct  | tcagcatcat  | gccagaaaag  | atctctaata | tccttaataa  | tatgtgaacg | 62820 |
| gctgttcacc  | atgtgtctagg | cgctgtgctg  | tagtatctca | tttgatccca  | gtgacaacct | 62880 |
| tatgggctag  | gaacacattt  | cagagggagg  | acacggatgc | ccagagactt  | caaacacctt | 62940 |
| tcttaagggtc | acacagctgt  | agactattgg  | agcaggagtt | cacatcaagc  | ctccctaaat | 63000 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| cccagacccc  | tgtttgtccc  | tctactctgc  | aaaatttcac  | atttgtttgt  | ttttagctta  | 63060 |
| taccacctct  | ctgaagggaag | ggttctttca  | gttcctatct  | gccgtataat  | gaggcatagg  | 63120 |
| tttttgttgt  | cctgtgaatc  | cttgggtgcca | gtttcagtag  | gaagaaagca  | agcctgactt  | 63180 |
| tagtagtaga  | aatacgagga  | gatagggcca  | caccatttta  | tgtatttgac  | aagtcgacgt  | 63240 |
| catagcatta  | aatccacagt  | actttttattt | ttatctttat  | tttcttattt  | tactttaagt  | 63300 |
| tctgggatac  | atgtacagaa  | catgtagatt  | tgttacatag  | gtatacatgt  | gccgtgggtg  | 63360 |
| tttgccgcac  | ctatcagcct  | gttatctagg  | ttttaagccc  | tgcattgcatt | tggtattttgt | 63420 |
| cctaattgctc | tccctccctc  | tgtccccac   | ccctgacagg  | ccccagtgtg  | tgatgttccc  | 63480 |
| ttccctgtgt  | ccatgtgttc  | tcatggttca  | actcccactt  | aagagtgagg  | acatggagta  | 63540 |
| tttggttttc  | tgttccctgt  | ttggtttgct  | gagactgatg  | gcttccagct  | tcatccgtgt  | 63600 |
| ccctgcaaa   | gatatgaact  | cattttatttt | atggccgcac  | agtattccat  | ggtgtatatg  | 63660 |
| tgccatgttt  | tctttatcca  | gtctatcatt  | gatgggcatt  | tggtttgggt  | ctaagtcttc  | 63720 |
| gctactgtaa  | atagtgtctc  | aataaacatc  | gtacgttaat  | atgtatctct  | gacaggggct  | 63780 |
| tttaaaaaat  | ataaggacaa  | tggtgttatc  | acatccaaca  | aaatttagta  | atccctaatt  | 63840 |
| ttatctaata  | tcaagtctaa  | aaagtatcat  | tttggacttg  | atttgtttga  | atcagggagg  | 63900 |
| atgaggtgtt  | tttaacatgt  | agcattaact  | aataaaaccc  | cttcagtact  | cattccatta  | 63960 |
| atggctttgc  | catattattt  | atcagcaact  | tcttcgagag  | atgcaacaga  | tgccagcccg  | 64020 |
| tccctttgccc | tctgtaaatg  | tgccttgga   | aacagatgag  | gagcctcctg  | atcttatttg  | 64080 |
| ggggagtaga  | aaggtagaaa  | tgtgactcag  | aagtccttat  | aacttgactt  | tttaaaactt  | 64140 |
| aggctcctaa  | gtctgggaaa  | ccagagagag  | caaaagccct  | attccattac  | ctcctttttt  | 64200 |
| tccctctctt  | ctttttaaca  | ataagctaaa  | acctgaagtt  | gctgggtact  | cttttgcctt  | 64260 |
| tgtttttcaa  | tctttgtttc  | atgtcgtgta  | gcagtcctgt  | gttcagggga  | ttttaaaatc  | 64320 |
| taaatgattt  | gtctgctttt  | cttttaagta  | ttctgtggca  | tacaaaactt  | ttatttgaaa  | 64380 |
| agagctacat  | tgaacatcga  | atcattgttt  | ttttgtttgt  | ttgtttgttt  | gagacagagt  | 64440 |
| cttactgtgt  | cacacaggct  | ggagcgcagt  | ggcatgatct  | cagctcactg  | caatctacct  | 64500 |
| ccggggttca  | agcgattctc  | ttgccccagc  | ctcctgagtt  | gctgggatta  | ccgatgtgcg  | 64560 |
| caaccacgccc | cagctgattt  | ttgtattttt  | agtggagatg  | gggtttcacc  | atgttggcca  | 64620 |
| ggctggctc   | gaactcctga  | cttcaagcaa  | tccaccacc   | tggcctccc   | aaagtgcctg  | 64680 |
| gattacaggt  | gtgaaccacc  | gcaccacgccc | tatttttctt  | tatatatttg  | ggaatttttt  | 64740 |
| aagcagcaaa  | atgataatac  | actgtgtatt  | aatactctgt  | ggaaaatatt  | tatcaccccc  | 64800 |
| ctccccaaaa  | tgatgaagaa  | aatggtaaga  | ctttctcctt  | gggcaaaatg  | atggaatatt  | 64860 |
| taagatacgc  | tggagaaata  | gctatgtatc  | ttgaataaaa  | catacttact  | taaaatgtta  | 64920 |
| ctgagctcac  | aggaagttag  | taaaatctcc  | agggaaatact | ctccttccct  | accctaaaaa  | 64980 |
| gaacaaacca  | taaacaaaaa  | ccagaaccat  | gatttgcctt  | aagtggttcc  | agatggtaaa  | 65040 |
| cacaacgccc  | acctgaatag  | atggaaatcc  | tctctataga  | aaaacatctt  | caattcagcg  | 65100 |
| ctgttgaagt  | cccacagatc  | aagatgaagt  | aattaaagat  | caccagggac  | aagaggcaag  | 65160 |
| ccgccataaa  | taagagttag  | agaaacaaca  | gatgattgat  | gtggacagga  | ctccaaatgt  | 65220 |
| tagaattgaa  | atctaaagaa  | tataatataa  | ctgtgtatga  | aatgtttaaa  | gaaataaaaag | 65280 |
| atgaaatcat  | aaagatgaga  | gtcaagtatg  | gccaaagcaga | tctgaagaag  | aaccaaattg  | 65340 |
| aacttctaga  | aatgaaaaat  | atacttgttg  | aaatttttta  | aataaataaa  | ctcaatttac  | 65400 |
| atgttaaaca  | gcatatcaga  | cataagagaa  | tttacaaggt  | ggaagatgaa  | tctgaaaaaa  | 65460 |
| gtacgcagaa  | tgcagctcag  | agagacaagg  | agatagaaag  | tacaatagat  | actaagaata  | 65520 |
| tggaggatag  | aatgagaaga  | tccaactcat  | atataatttc  | agaatcccag  | aacaaaggag  | 65580 |
| aggcaatatt  | cagaaatgat  | agctttccag  | aactgatgga  | aaacatgaat  | atacagattc  | 65640 |
| cagaagcaaa  | acatattgta  | agcatgatta  | aagaaagaaa  | gggaaggcag  | gaaggaaaga  | 65700 |
| aaacaagtgt  | gatttctcat  | aatgaaattg  | tataactcca  | aagacaaaga  | gaacatctta  | 65760 |
| aaatcagcca  | gagaaaaagt  | atcacataca  | acgtggagaa  | actgaacggt  | agtgactgtt  | 65820 |
| tgatggaagt  | ataaagtgat  | ggagccactg  | tggaaaacag  | taccactatt  | ccccccagaa  | 65880 |
| ttaaaagttag | aattatatat  | gatctggcaa  | tctgggttta  | tacatagaag  | aactgaaagc  | 65940 |
| agagtctcaa  | agacatattt  | gaacatcagt  | gttcttggca  | gcattattca  | cagtagccaa  | 66000 |
| aaagtggaag  | caggcccaag  | tgtccatgga  | tggataaatg  | gataagcaca  | acgtgggtcta | 66060 |
| tgcattccaat | ggaatgttat  | ccagccttcc  | aaaggaagca  | aattctgacc  | catgctacaa  | 66120 |
| tgtgatgaac  | cttgaagaca  | tactaagtga  | aatactccag  | taacaaaaaa  | ataaacattg  | 66180 |
| tatgattaca  | cttttaagag  | atacacttgt  | ggtagtcaag  | ctcataaaga  | cagggaagtag | 66240 |
| aacagtgggt  | cacaggggct  | gggagaagg   | gaaaatgggg  | agttagtgtt  | taatgggtac  | 66300 |
| gaagtttgag  | ttttacaaga  | tgaaaagagt  | tctggagatg  | gatagtaatg  | atgggtgcac  | 66360 |
| aacaatgtga  | atataacttaa | taccactgaa  | ctgtacattt  | ttaaatgggt  | aagatggtaa  | 66420 |
| atgttatgtg  | tattttatca  | caatttttgt  | taaaaatggg  | aaaagaggcc  | gggtgtgggt  | 66480 |

26/154

|            |             |             |             |            |             |       |
|------------|-------------|-------------|-------------|------------|-------------|-------|
| gctaacacct | gtaatccgag  | cactttggga  | ggccgaggcg  | ggcggatcac | ttgaggtcag  | 66540 |
| gagttcaaga | ccagcctggc  | caacatggtg  | aaaccccctc  | tctactaaaa | atacagaaat  | 66600 |
| tagccaggca | tgatggcaca  | catctgtaat  | cccagctact  | tgggaggctg | aggcatgaga  | 66660 |
| atcatttgaa | cctagaaggc  | agagttcaca  | gtgagctgag  | attgcaccac | tgtactccag  | 66720 |
| cctggggcag | agaacaagac  | tctgtcgcaa  | aaaaaaaaaa  | aaaaaaaaaa | aaaagggaaa  | 66780 |
| agatcacctg | caaaggaatg  | acaagctgac  | cgctaacttc  | tcagaagtca | gaagagagtg  | 66840 |
| agagaatgtc | tttgaagtac  | tgaaagaaac  | ctccccacct  | agaatagtgt | gcccagtaag  | 66900 |
| caccttccga | gaacaaggat  | gaaataaaaga | tatcttcaaa  | ttagaataaa | aatttgagaat | 66960 |
| ttgccaccag | cagacctgca  | caaaagaaat  | gtctgaagga  | tgtgctaaaa | tgaaggaaaa  | 67020 |
| ttacccccag | gaggatagac  | taaagtgtaa  | gaaagaatag  | tgagtaaata | aaagtataat  | 67080 |
| catatataga | aatctaaaca  | aacattatcc  | tggtaaaaata | atatttgtgg | gttaaaaaat  | 67140 |
| agacaagcct | aaagtattga  | acaacatgat  | atatgtcagc  | agcatatgac | acagttagtt  | 67200 |
| gctccttaaa | acattttctt  | gatttttctt  | caggacacca  | cacacttgcc | ttctagctgt  | 67260 |
| tcttgctcag | tgggtctttg  | tcagtctctt  | tcactttttg  | acagtgggtg | gcgtcaggcc  | 67320 |
| ctgcctttgt | cccttctctg  | tgtctttttac | accacttctt  | tcagtgatct | tcactatcac  | 67380 |
| aggtttaagt | accatgatcc  | tctcacctgg  | actgctccca  | tgagcctcca | acacaaacat  | 67440 |
| cagttgccta | ctcagtgtcc  | ccacttggat  | atcttaacat  | ggccatggcc | aaactgagct  | 67500 |
| ttggatgttt | attttctcaga | cttgctcctc  | tgacataacct | tccgcctttc | agcacataac  | 67560 |
| acctccatca | ccctagtgtg  | tcaggcccaa  | atcgtggaat  | tgttcttgac | tcattcttga  | 67620 |
| tgtcatgttc | tcattggaca  | gccagtctat  | cagcaaatct  | tgactccgta | ttcaaaaatg  | 67680 |
| tttcccagcc | actgtgtccc  | atctctgcca  | ctattcccac  | ttgatctagg | ccaccgctgt  | 67740 |
| ctctcacctg | gatttcttga  | gttacttcat  | aactgatctc  | cctgcctctg | ccattcacc   | 67800 |
| atctagtcag | ttccttgtgc  | cacagcagag  | tggtagtgga  | aattattttc | ttcagccaaa  | 67860 |
| ctgaagaaac | tgatgacatt  | tgtgggtata  | gtcagcagaa  | ttctaaaggt | ggccctcaag  | 67920 |
| atttccatcc | cacgtttatt  | cagtcacaca  | ctactctagg  | tactgctaga | gggattttgc  | 67980 |
| agatgtaatg | aaacaaggga  | gttgacttta  | agatacagaa  | agtttctctg | gctggtggca  | 68040 |
| gaaagaggaa | gccagagaga  | ttggaaacat  | gagaaggact  | tgacacacca | tccctgggtt  | 68100 |
| agagatggag | tggccacaga  | agtttttttg  | tttttgtgag  | acagagtctg | gctgtgtcac  | 68160 |
| ccacactgga | gtgcagtggc  | atgatcatgg  | ctcactgtag  | cctcaaactc | ttgggctcaa  | 68220 |
| gggactcctc | ctcctgggtc  | tccc aaagtg | ctgggattac  | aggtgtgagc | gactccgtaa  | 68280 |
| ggcctgaatg | catttcaaaag | cagagtcctc  | cccagagtct  | ccaactaaga | gtccccatcca | 68340 |
| gctgacaact | tgacttcagc  | ctgagagacc  | tcaggagata  | gctgataatt | ccccagaact  | 68400 |
| catggaagac | atgaacagac  | ctgtctctct  | cggcagagaa  | ccctacagag | cttgcctggga | 68460 |
| cttctgacct | ctaaagctgt  | gagataataa  | gtgggtgttg  | ttttaagcaa | ctaagtttgc  | 68520 |
| agtagtttgt | tacactaata  | taagctgggt  | gcggtaggct  | acacctttta | tcccagcact  | 68580 |
| ttgggaggct | gaggtggggag | gatcgcttga  | gcccaggagt  | tcaagacctg | cctggacaac  | 68640 |
| atagtggagc | cccattttcta | ttcttttaaaa | aaaaaaagga  | aaggaaagaa | aaagagaaga  | 68700 |
| aactaatata | ggggacaaaat | ctcatctaaa  | ggctgcaaca  | gaagatgaaa | acatattttc  | 68760 |
| cagagccaaa | tgcattccaga | gagtagacaa  | agaactctact | tggtagccca | accagctttt  | 68820 |
| tacaagcttt | tgcattccaga | aggcttctct  | acttagtgct  | taaaggacag | acaaaataga  | 68880 |
| atgtttttta | agtctccatg  | atatctgcat  | ctgtgtttga  | gccaggtagt | cagaactatc  | 68940 |
| tgacaaatct | accaagcact  | caccactgcc  | attcctgtga  | cctacagtta | tgaattagaa  | 69000 |
| tttcaaactg | agccaaagta  | cagcatccag  | aagaaaccca  | cttggcgtgg | ccccctccct  | 69060 |
| acagttcaag | cagccctctg  | gaaggggact  | gcgagacctc  | agtggaaagg | agcaagcagt  | 69120 |
| accacccgtg | tcactgggtg  | tcagagagta  | ctgttagctc  | tctcctgccc | tttttcatta  | 69180 |
| tcttaaaaa  | tgcaatgtac  | ctttttttata | tacctgaaat  | tgatggttct | tataagcaca  | 69240 |
| aatgaccggt | aagttgtccg  | tatacattaa  | agcttgagta  | tcacttaagg | taatttttgt  | 69300 |
| ttcttctgcc | agaactgttc  | caaaccattt  | gcactggagc  | cgtgttttgg | caacaaagcc  | 69360 |
| gctgtcctct | ctgtgtttgt  | gaggctccct  | cgaggcctgg  | gtggcatccc | tcctcctggg  | 69420 |
| cagtcagggt | agtagatgag  | gtccagcgaa  | agacaccttc  | taagcatgtg | agggagctaa  | 69480 |
| gcatgggata | cttccctttc  | ctagagaaca  | aggttataaa  | ggtgataacc | aacagtgcct  | 69540 |
| gcccattgca | tggcaaggat  | tcagttagat  | aacctatttt  | cacctatatg | cccaaaactg  | 69600 |
| ttttcccag  | gtcatcgctc  | cttccctggg  | ggacgtccc   | ggacctcctc | accgttctgg  | 69660 |
| cctccccagc | ctctcagtg   | tcgtggatgc  | ctctgaccac  | accttccttt | gctgtcggag  | 69720 |
| tttctgggtc | ccctttgccc  | ctgctgctct  | ttccctgtct  | cttttgcctt | gtcctgcccc  | 69780 |
| tgaatctggg | tgtgtccctt  | tcaccctgca  | tgttcttcc   | cagttgcctt | ctccactcca  | 69840 |
| caacttccag | ccatctcccc  | tgggctgggt  | tttcttcaat  | tggatatttt | agccctcctg  | 69900 |
| catttctttt | gctttttcac  | tgtttttaaat | ataaggcaca  | aacacaacga | taacataaag  | 69960 |

27/154

|             |             |             |             |            |             |       |
|-------------|-------------|-------------|-------------|------------|-------------|-------|
| gaaacgtctg  | aaaatccacg  | cacagttctg  | ctatcccaca  | accaccaat  | tttcattggt  | 70020 |
| ctgtgttctt  | tccagccctg  | gccagcaatt  | tagttgaaat  | gattgtatag | atctaatttt  | 70080 |
| tattctgctt  | ttttcatata  | actatcttaa  | acatttttat  | agcctactac | ataatatttt  | 70140 |
| accaatgcat  | gtgacacttc  | tccatcaaat  | atatgaatca  | tgattatata | aaacttctatt | 70200 |
| ttgaagacag  | aaaatatgga  | atacagtttt  | gaaagtatct  | cgcttaatta | cagagtgcct  | 70260 |
| tggaggatgc  | ctcaatgtga  | tcatitttaga | caactttcag  | ttgagatgat | tctggagtca  | 70320 |
| gggttacctt  | agccagagag  | caccctgtgt  | ctggtggaaa  | ggctgcccc  | ggatactggt  | 70380 |
| ccaggccctg  | agctccactg  | tcccttcagt  | gcaccctcag  | accagaaga  | tcctccagcc  | 70440 |
| cagcttttaga | atggtttagga | tctatagtat  | cttcacgagg  | gaggcctttc | gacatggctc  | 70500 |
| ggccaggatt  | cattgtcttc  | atctcaccta  | gggattcctg  | tgagttgagc | ccctaactcg  | 70560 |
| gccttgctcg  | tgttgaggac  | aggacaactc  | tctcgtctgt  | cactcaagtg | cagtttgtgt  | 70620 |
| ctcctgccac  | cgcccagccc  | agcaaacagt  | ggctgggtgtg | ccctgagggc | aggggtggcg  | 70680 |
| ctggggcggc  | aaggtgacag  | gtgtgcagcc  | ctgtagcctt  | tcctgctcca | gcctttgcct  | 70740 |
| gacatgcccc  | aaccacctga  | aacaccta    | aaaggtctgc  | tcaatgcttg | agaccacccc  | 70800 |
| aacagtgggg  | ccgtccttgc  | tgcctcttct  | gtacccctcg  | ttttaagtgg | tgcccttttc  | 70860 |
| taggctccca  | gagcattccc  | cactgtgtgc  | caggcactta  | cagctgtcta | ccccacccc   | 70920 |
| cactcccctt  | cacaggctca  | acaaggcaag  | taacacgctc  | cactaacaca | actcctggca  | 70980 |
| ggtctcagca  | gagctccgag  | cctgcgctgg  | ccatttccca  | cgtatattag | gaatccagct  | 71040 |
| cccccagaac  | cgtgccttcc  | ctaattggatg | tcagattact  | ccgttatcaa | tacaggaagg  | 71100 |
| aaaatgggat  | tgcaagtcct  | caaagcaaa   | tgggcggcgg  | ggcagtgatg | agcataaagg  | 71160 |
| gtagttgcta  | ttgatcgtgg  | gttgctgcgc  | gccagatact  | aagactttat | gcattgccag  | 71220 |
| tactaagact  | ttgtcagttc  | actcattgat  | tcttcacact  | agccctgtac | tactattatc  | 71280 |
| cccattttac  | aaatgaggaa  | actgaggtgc  | agggaggtta  | aagaaccttg | cttattccaa  | 71340 |
| agcctggccc  | cagagcctgc  | cctgcctgta  | tcgtggatg   | gaatggaagt | gactaatgct  | 71400 |
| gactgagtga  | taggggtgaa  | tctgaatctc  | gtcataaagc  | gtgctgctat | ttctcggtc   | 71460 |
| cttttgagg   | aaggcaggag  | aaattgggca  | cttagcacac  | gtctacaatc | tcattttatat | 71520 |
| gttcccacca  | ccaagcagag  | cactgtgctt  | ggttgatgtg  | ctgaaaagca | aaactgcacct | 71580 |
| tttgtgtaaa  | aagtagggaa  | aggagaagaa  | tgtgtatttg  | tatttgcata | agtataccct  | 71640 |
| caaagaatac  | ataaaaatta  | gtgaaagcgt  | ttaccagat   | ggaggggaag | gcggagggag  | 71700 |
| ggtcagggtg  | gaaggaagac  | ttttccctgt  | gtatgtgtac  | ctttttgggg | gttttttgtt  | 71760 |
| ctgggttttt  | gggtttgttg  | ttttgcacca  | tgggattgaa  | aaataaatgt | ttaaattatt  | 71820 |
| ttttaaaaca  | aaataaaagc  | cagctgctcc  | cagcatatgt  | tctctctgtg | gttctcccaa  | 71880 |
| ggtcttgctg  | tggccagcgc  | cctgggtggac | atttctcagc  | agatgccgat | agtgtacaag  | 71940 |
| gagaagtcag  | gagccgtgag  | aaaccggaag  | cagcagcccc  | ctgcacagcc | tgggacctgc  | 72000 |
| atctgatgct  | ggggccagg   | actctcccac  | gcacgagcta  | gtgagtggca | caccagagcc  | 72060 |
| atctgcagg   | aagggcgtgg  | cggggaaatg  | gctgtgcgg   | gcgggacgga | agactggaaa  | 72120 |
| ccctcaaagc  | atctgactca  | cctgcatgat  | cacaagcttt  | ctttgacgg  | ttctcccatc  | 72180 |
| cgtgttccag  | catctaacct  | tttacttttg  | cataggaaat  | acttgattta | attacaggct  | 72240 |
| cagggatgag  | ctgtaggttg  | ctggaggagg  | ccagtgtaga  | gccagtgaga | gaactaggaa  | 72300 |
| tgacactcag  | gttcactgtg  | gaaaactgtt  | cttgggactg  | tctcaactgt | gcaaaaaaca  | 72360 |
| aaagatggag  | tgtttacaag  | tagacattcg  | tcacagttg   | ttcttgaaca | tgggtctttta | 72420 |
| aaaactagtc  | agatgaatta  | acttgttttc  | atctgaagcc  | tgctatcttt | tttaaaagat  | 72480 |
| gtgctattta  | ttcttgcaag  | atttaggcaa  | ttatctctct  | tccaggaggt | accttttttt  | 72540 |
| ctagttgaga  | attaataatg  | gtccatctct  | tttgatcata  | tcaagctagg | atagaagggg  | 72600 |
| ggctatttta  | aatgtcaagg  | tcagcagtg   | tactttgaat  | gtaaaactgg | ataataggta  | 72660 |
| gttttctata  | gtaacttgat  | taattttagt  | ttaatccatt  | tgaaactctc | tcttcccttc  | 72720 |
| tctctgctg   | tcctctcct   | tctccatctc  | accctccctc  | tctcacacat | acacacacaa  | 72780 |
| acacatacac  | acaacactaa  | gtgcctagac  | tttaaataga  | tctagcaatt | ggaaagttag  | 72840 |
| taagcctaag  | tttttacata  | attgcattcc  | tacattcttg  | taaaatttaa | atagctacca  | 72900 |
| ttggcaatct  | gctttttttc  | taaaatctga  | tttgacgcca  | ggaaagaatt | ttctcaccca  | 72960 |
| aggaacattt  | gatctagcag  | cagggatgag  | aggaaagcag  | aatgaatga  | actgtgaaag  | 73020 |
| ctcctgtttt  | tattatcaaa  | aaggacactg  | tcaagaaggc  | gccccctgcc | cccacccccg  | 73080 |
| tgtcacccta  | ggcctgataa  | gcgatcagag  | gaaaggactc  | attcatgtca | cgcttccttg  | 73140 |
| agcagaaaag  | agcactgaga  | gcacttggga  | cccctggatc  | agagagcatc | tgtgtgtcct  | 73200 |
| gcagcctcct  | ctgaacttgt  | ggttcattct  | caggctgggg  | tggactcaga | tgccaggaaa  | 73260 |
| gggacagcct  | cccatttgtca | ggcagaagct  | gccccaaagc  | tggagaagga | cttgttttg   | 73320 |
| ctctttcccc  | caggaggggc  | tcgacccacc  | caccctccct  | ctcagaccaa | ggtggtggct  | 73380 |
| gtgaggaggg  | cagcaaatgc  | tgacaaggat  | gaaaagcaca  | tggaaaaaaa | tggacgagga  | 73440 |

|            |             |             |             |             |             |       |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| gggaaaactc | tgccaaatgg  | aaaatgacca  | aattttaagag | gggaggacag  | tccctctgctc | 73500 |
| ctctcccaga | gggcactgct  | tggaattgt   | gttttcccca  | tttatgggtgc | tctgtattct  | 73560 |
| ggcattatgc | agcagcctcc  | cagaagctct  | cttctgcttc  | aaaacctggg  | atctctggca  | 73620 |
| ttaccctatt | gggatggacc  | gctggacagc  | aatgctcgag  | tttgtgaatt  | tggagagata  | 73680 |
| ctcaaaagag | ctaaaactgc  | agcattttac  | ctttaaatgc  | agtgcctaga  | gagagagtat  | 73740 |
| tgtctcttcc | ccaacactaa  | ccccactccc  | atgaagaatt  | gcctggaaag  | atgttttcaa  | 73800 |
| ggaatttgaa | ccataaaaca  | ctatctgatg  | cacagaacac  | ctctactttg  | agactcacct  | 73860 |
| ctcataaagc | ttctttttca  | cattactggt  | aaagaccaga  | cgttctagaa  | aagaccctc   | 73920 |
| ctctcatgag | ctcccccatc  | cctgctacag  | aacacagcac  | ccatggcgcc  | tgcagtggac  | 73980 |
| tggcccttta | attcccacag  | gccccccag   | caaggccaaa  | gggaggcccc  | tgggtattgt  | 74040 |
| cctcctacaa | ggaagatcct  | ctttgtttgt  | tcaaaggacc  | agttttcccta | ggccaaagaa  | 74100 |
| gtctcttccc | catgttagtc  | ctatgccttg  | aaatatcatg  | caccatgacc  | cacagccatc  | 74160 |
| tggttatgtc | ttattttttt  | cctaaaagat  | aatgtttatt  | tttaaaaagg  | aaggaaggag  | 74220 |
| caagtgaagt | ttcattctgc  | tccagcgtg   | gggaagccgc  | tgaatccacc  | tgttctcct   | 74280 |
| ttgcaaccga | cagcaaacag  | ctttctccgg  | cctcagggca  | gaaaaaggga  | atggcaggga  | 74340 |
| gtaagaggcg | ctgggctcgg  | agcctgtttc  | caagaaggaa  | ttggttgtca  | tctggcagt   | 74400 |
| ttgcgcgtca | caagagagcc  | tgtatataaa  | ttaaaatagt  | caagacaaca  | ctgaccttgc  | 74460 |
| acttgtacat | aactatacag  | tagtgtccag  | aatgttcaga  | cattcggagt  | gtacataaaa  | 74520 |
| cagaaaaaat | cttcatgtat  | ttttattaaa  | tataacaatg  | tctgagtttc  | acctaagatg  | 74580 |
| tttttgtgcc | atatgctgga  | tatccaggtt  | ctcgccaggc  | cccatacat   | gaataacaaa  | 74640 |
| cccaagaaac | gcatcccat   | tgtgtgatgt  | gttcagatgc  | atctggcacc  | aattaggtat  | 74700 |
| ttcttaaaac | aggactcatc  | tgtcagagt   | cacatgaaaa  | atcaggcagg  | gaatcgaaac  | 74760 |
| gacagcgctg | gaggagactc  | aggaagcaga  | ggcgtccctg  | ccgctgcct   | tggccctgca  | 74820 |
| agcacatcat | gaccctttct  | ggcagcctct  | tgggtgctctg | ggtagtgagg  | gatgaccagt  | 74880 |
| cttgtcctga | gaaatgtttc  | tcttagtctt  | taagttcaaa  | gactaacctg  | tagcaatcag  | 74940 |
| actttccaaa | aggggggttct | ccattttttg  | tagttttgtc  | taaattttta  | atgaccattt  | 75000 |
| cctggaatca | gtttattata  | ctgaaaactg  | ggggtgggag  | tagggagcta  | gtttgttgat  | 75060 |
| aaatagttcc | catttccccg  | tggagaattt  | gacataccct  | ggactcctgt  | gtgcctcctg  | 75120 |
| ccatccctgc | acacagcctg  | gggagaagcc  | tgtgcctccc  | cgtgtggaga  | gaaggcaacc  | 75180 |
| ccagatcccc | tgagctaacc  | cggaggaag   | cgactcctgg  | acagaagact  | gtcagcagaa  | 75240 |
| ggaaagtact | ggactaccgg  | tgggttaagtc | ctgccattca  | agactggaga  | cacctgggaa  | 75300 |
| ataaaaagag | cagggcactg  | ctgggtgggaa | gaggcatttt  | accttccagt  | gcaaatcctg  | 75360 |
| ctcctttgat | ttaatggggg  | gtactggggc  | caggggctga  | ttcacttcc   | tgggagatgg  | 75420 |
| tgggtgtttt | atgaacatct  | ttgatccttc  | catttcattt  | attcatccat  | ccattcaaca  | 75480 |
| agtatttgct | aaacactaac  | ttaagctaatt | gctagggtag  | tgactgagat  | gtaaaaatag  | 75540 |
| attttagaat | taaaacaaaa  | tccaagtcc   | cacaccctg   | tcatcccagg  | agatctttcc  | 75600 |
| ttgtgggtgt | ttctgtgaga  | attggccatc  | ctgaggacac  | agccaggacg  | gcagaggcct  | 75660 |
| cctggcctca | gggcactgcc  | tgccctacct  | ctgaaatgtt  | tacccatttg  | accaaacttg  | 75720 |
| gctccagcca | tgcggtggt   | ttctagatag  | ccaggccac   | caagagatat  | tggcccttga  | 75780 |
| tgagagtcaa | acacgctgcc  | tacaaggaga  | tgttttgaaa  | tggagaggaa  | aattggcacc  | 75840 |
| tcacttttta | aaggcagtaa  | tgggaattgat | tttcagtaac  | tgaatttggt  | cacaaaacat  | 75900 |
| tctaacaact | agtgaagcct  | gtttcggtga  | actaattaat  | tctggctctg  | gaaatgtttt  | 75960 |
| tgttttatag | ttatttacga  | tttcggttgt  | ttggattcaa  | gcttagtttg  | ttaatatgta  | 76020 |
| taatttagca | tctattacac  | tcattgtaaat | atggagtaag  | tattgtaaac  | tatttcattg  | 76080 |
| cggggattgt | gggtgttata  | catacattta  | ggactgcaat  | tttttggtat  | ttttttgtat  | 76140 |
| tgtaaaataa | cagctaattt  | aagcaggaac  | aagagaacta  | agggaggctc  | gtgcatttta  | 76200 |
| aacacaaatg | tgaagaactt  | gtatataaac  | aaaagtaaat  | actataatac  | aaacttcctt  | 76260 |
| ctgaaataaa | agttagatctg | gtaaaaatgt  | ggcttttgtt  | ctgagtgttt  | cattttgatt  | 76320 |
| ttgcattggt | ttgcctatat  | ttacattttg  | gtcattagaa  | ccttaaggga  | aaaaaaaaag  | 76380 |
| cgtttaccag | ttttcagact  | gcgtcagagg  | gagcacttga  | cagttaacgg  | aggtgagggt  | 76440 |
| gaaaaccctt | ggcaggtagc  | tcagacagcc  | ctccaggcag  | cttgtgagcc  | acagaccttc  | 76500 |
| cacacagcct | ttgtgtgcct  | acagggcagc  | tgggcccgtg  | ggggcaggga  | cagatctgag  | 76560 |
| cagtgaggta | tggacagccg  | tctcactgct  | ccactaacca  | caccactgtc  | tccagtgggt  | 76620 |
| ctgttatccc | tccactttgc  | gctctcctca  | caggctcctg  | aatgtcctct  | aggttgtctc  | 76680 |
| caaactgcta | aatctcaggt  | gtgctgacag  | gtctcagccg  | tggccagtgt  | tttctcggc   | 76740 |
| cttgaagttt | tgaggaagtt  | actcccattg  | gcaaggcctt  | gtttgggggc  | ctcctgtgtg  | 76800 |
| tccctgcccc | ttttattttg  | ctcctgcctt  | cagaatgggt  | cagagtcatg  | tatttcagt   | 76860 |
| cagagtgcag | attcccagtg  | ctagcagttc  | cttcagaag   | gatgaggta   | ttgaaagaag  | 76920 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gaggaaacac  | aagacgggtca | ctcctgtagg  | ggaagatggc  | agggtagtag  | ggccaggggc  | 76980 |
| tgaggatgaag | gcaggagcct  | ggaggaggcc  | tgctaccag   | agctcaggtc  | tggttaggg   | 77040 |
| agggaaatgg  | ctggactagt  | ggctgcaaac  | aagttagaaa  | ggaccttgaa  | actcttcag   | 77100 |
| ggagggttat  | gagagccacc  | aacagtaaca  | ggatagaaac  | cctgtgattt  | ggtggtactg  | 77160 |
| tggcagagac  | aatgctgtgt  | ttaccaaacc  | tactggggcc  | acagctcatc  | atttcccagc  | 77220 |
| ctccctgcag  | ttaggttggg  | accaagctaa  | gccatggcca  | aggaaagggt  | ggaagaagtg  | 77280 |
| atgggcacca  | attccaaaacg | tggcacacaa  | atgctccac   | aggggactct  | cctacggatc  | 77340 |
| acatgttaaa  | gactggaggg  | actgtgggtc  | cctgagtcac  | cccttagagg  | ccggccgctc  | 77400 |
| gatcagtaag  | agctgtgtgt  | gcctttgcac  | cactggaaag  | caaactcact  | ctgctgatcc  | 77460 |
| actgagattt  | ggggctttgt  | tacactagtt  | agcttttctt  | attctcacta  | atatacaaat  | 77520 |
| cagcactttg  | aagtgcagtg  | cacatgtaac  | caaaccctta  | agtacatggc  | actggctcag  | 77580 |
| cgtgtggcca  | ggtggcgagc  | agtataaga   | cagggtgcagc | tggcaggaca  | actgaggatc  | 77640 |
| ctgccaggcg  | gagcaaaaca  | tctggtaagg  | ccgtttcact  | cgataactcg  | gaaaacaaac  | 77700 |
| tccatgtctg  | ctgacttttt  | gacttgaggg  | gaataaaca   | caacaacaac  | aacaaatatg  | 77760 |
| tgttgattgt  | tactggctgc  | ttttggcgag  | gcacacagag  | aaaaagttta  | cttcaagaaa  | 77820 |
| gaattagcca  | gtttacaagc  | agaagggaaa  | aagatggaaa  | gatgttaaac  | agggccttgc  | 77880 |
| aattctggaa  | aagccaactg  | cttctagtct  | ccaaccagta  | agagggtgcag | gtgagcaaca  | 77940 |
| aaggccccat  | gagacttagt  | tgaacacagt  | gactgggctc  | catggtgacc  | atccaatgaa  | 78000 |
| gggtgtgggtc | ttcccactgg  | agcatggaaa  | ctgcaaagcc  | accagcttga  | ggggaggggc  | 78060 |
| cgtgggttaag | gaagcaaaag  | aaaagctgac  | atgagaactg  | tttccaggaa  | caaaccatga  | 78120 |
| atgtagtcac  | tgaccctgga  | cctgcctgga  | aataaataaa  | gatgtagtaa  | atttaagaaa  | 78180 |
| ccattagcaa  | agccttaaaa  | aaaaaaaaaa  | aaaaaaaaaa  | aaacggtggc  | aattgaggaa  | 78240 |
| gtcttaaatc  | caccttcagg  | caggaaaatg  | gttgtgaaag  | ctgtgtccag  | cgagggtctt  | 78300 |
| acaagctcag  | atgtggggcc  | aagagcagaa  | ctgggcctga  | gcgaggactc  | ccccagccc   | 78360 |
| aggacagggc  | ccagcaggat  | tgcaaaaaca  | ccgcagagca  | gcaattgctg  | catctgcacc  | 78420 |
| cgctccttcc  | tgggtaggag  | ggcctaggag  | ggtgtcctgc  | ctgttctcat  | cactacctgc  | 78480 |
| tgggtgggag  | gtggattaag  | ctgtcttgtt  | aattcaggat  | tgtggacttc  | aaagagccac  | 78540 |
| ctgaggggag  | ctgcactacc  | tggaggtctg  | ggcgggggtg  | gacctcgtat  | tgtccccta   | 78600 |
| ggagagggtg  | cctgtgaccc  | atgtaggaga  | atgagcacaa  | aactgggttt  | ggtgatcaga  | 78660 |
| atagcaaaat  | gggccagatg  | cttaccacac  | atctcctctt  | cctgagtcca  | cagcttcgac  | 78720 |
| tgcattggct  | gggcccgtgg  | catctacca   | gatgccttgg  | gaccagccct  | caccaatgat  | 78780 |
| gtgtcatttc  | aaggccattc  | acagtgcag   | tgtctcctcc  | actgtctcca  | tcaaccttga  | 78840 |
| aggcttatgg  | tgaagatgg   | aatacaacag  | gattgaagga  | gactgggtcc  | ctgaatgact  | 78900 |
| ttgtggaaca  | gagcactata  | ggccaaaggt  | agtagaatta  | ggtatcttgt  | tttatgctgc  | 78960 |
| ttaagtgaat  | taattaattc  | tctagattta  | ttcacctagt  | tattcaacaa  | atthttattg  | 79020 |
| agcacatcag  | ggctaggggc  | ttgtgtcagg  | atcagagatt  | ttgcaatgaa  | caaaacagaa  | 79080 |
| aactccatgt  | cctccagggg  | agttgataat  | aaaatgtcag  | ataagctctg  | tgaagaaagg  | 79140 |
| taaaagtagg  | gttaggggta  | gaaagtgtctg | gggggattgc  | tacttaacat  | agtggtggcca | 79200 |
| ggggaagcca  | ttctgtctaa  | gtgacagctg  | agcagaacga  | tgagtggagg  | gggcagggaag | 79260 |
| gaatgtggag  | gtctgggaga  | agagcattcc  | aaacagggat  | ggcagggtgc  | gaggccctga  | 79320 |
| gggggcagca  | tgcctggtag  | gcctgtctga  | tatttctgag  | aatgaaaaag  | tccattccca  | 79380 |
| gttgttttgc  | tggagaacct  | aaccagttta  | taaacacaaa  | gtacttgaga  | gctggtttta  | 79440 |
| tatacagcca  | gctttccagc  | agtgcactta  | ctgtgtacat  | caagggaaaa  | ctgaacttcg  | 79500 |
| ttttccttaa  | aacttatcat  | cagctgggtc  | tcattttgac  | aaattctgtc  | aacaacagca  | 79560 |
| gtgtcattcc  | tggcatctgt  | atgggtcacg  | tctgaacaga  | cacacgccct  | gcagccctgc  | 79620 |
| aggtaccagc  | tgtataacaa  | gaactccctt  | ccacctgtgt  | tcctggaaac  | aagaaagcca  | 79680 |
| ttagaccgga  | agatcccgat  | ggctatctca  | aatgtgtctg  | atggagttgc  | caggggccac  | 79740 |
| tggcatgccc  | tgtaaagcct  | tccttccacg  | tttggttcct  | gccccttgaa  | gactccattt  | 79800 |
| ctgagtttgt  | gtgtgtttta  | ctttctagt   | tgtgtcctca  | tcttaatttt  | tctctctctc  | 79860 |
| ttctgccttg  | aactgaagg   | tcgcttgggt  | gtggagagac  | agggccccag  | cagagcagct  | 79920 |
| tcccagagaca | tcctccgac   | cagggtctcc  | cagcagccc   | gcaaggcagg  | gctgtgcctt  | 79980 |
| tctgcttcag  | ctcacaagca  | tgccagggtc  | actggcaagc  | tgctgtctgg  | ttgagggact  | 80040 |
| gctcctaaag  | ccctgcacag  | cccctgtcct  | cctggccctc  | tggaaattcc  | acccccgtgt  | 80100 |
| ccacatttca  | tgcaaaaatg  | agctgggtct  | gtgagcatgg  | cccggcctga  | ctcgcttagt  | 80160 |
| gggcggttaag | tggtttccac  | ttcaacctgt  | cacctaatca  | ccgggctcca  | caccaggatg  | 80220 |
| gacattcatg  | agccgtgaag  | tttccagtaa  | taaatccaca  | gatgcttcca  | gcacctgcct  | 80280 |
| tttgcacatca | cctccactcc  | cagccacctg  | ccaggcaaca  | ggtaacagag  | accaggtcac  | 80340 |
| aggagggcag  | tgtgggggca  | ggactgcagt  | ctcccaaagc  | ccatgcacaa  | aaccgacagc  | 80400 |

|             |             |             |            |             |             |       |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| gccctggcag  | gacaaggagg  | ctgacattca  | gatgtggagg | aacaaggcat  | gacccattcc  | 80460 |
| tggtcatggg  | ggccacagct  | ggactcagcc  | ttgaggcttg | gccagactta  | acaccgtgta  | 80520 |
| taaaccagga  | ccttttttagg | tagagtaatg  | gaaaccaaac | tctaattgatc | ttagacagtg  | 80580 |
| ctattagtct  | cctggagctg  | ccagaacaaa  | ttaccacaac | ttcagtgtct  | tgaacaaga   | 80640 |
| gaaactgatt  | ctcacagttc  | tggagaccag  | aagtctgaaa | tgcaggtgtt  | gccagggctg  | 80700 |
| tagtctctgg  | agactctagg  | ggaatctgtg  | cctacctcct | ccagcttcca  | gtggctcctg  | 80760 |
| acattccttg  | gcttgtggct  | gcatcacccc  | aatccctatc | tctgtcttcc  | cctggctctt  | 80820 |
| tgctcaaaat  | gtctgtgttt  | agtttcctctg | tagacacctc | tgcatacctc  | tcataagatg  | 80880 |
| cagagggtgcg | acatacaggt  | gttgagagcc  | cacttagata | atccaggata  | agctcctctc  | 80940 |
| aagatctgta  | acttggtctgg | gtgcagtgtc  | tcacacctgt | aatcccagca  | ccttggggagg | 81000 |
| ccaaggcggg  | agggtcactt  | gaggtcagga  | gttggagaac | agcctgggca  | acatggggag  | 81060 |
| accctgtctc  | tactaaaaat  | agccaggcgt  | ggtggcacac | acctgtggtc  | ccagctactc  | 81120 |
| aggaggctga  | ggtaggagga  | ttgcttgagc  | ctgggagttt | gaggctgcag  | tgggctatga  | 81180 |
| ctgcaccact  | gaattccagc  | ttgggtgaca  | gagtggagct | gtctcaaaaa  | aaaaaaaaaac | 81240 |
| ataaaacata  | acttaaataca | catctcttgc  | cacagaaagt | aatactcttt  | tgcctacata  | 81300 |
| taaggtaata  | tttacaggat  | ccaggggtta  | ggatgtggac | atatcttttg  | gaccactgac  | 81360 |
| agccatgaag  | caatctcata  | attttcaaat  | aggttctgtc | ccttttatct  | ttccagtctt  | 81420 |
| ttggaaagca  | tatgcttata  | ttttcaatcc  | acaattctat | ttttatttga  | ggtcatttca  | 81480 |
| tttctgggtt  | ttatttttta  | ttgagacagg  | gtctcactct | gtcacaggct  | ggaatacact  | 81540 |
| agcacatga   | tgggtcactg  | cagccaactt  | ctgggtgaa  | gtgatcctcc  | agcctacagc  | 81600 |
| tcttgagttag | ctgggaactac | agacacacat  | caccatgcct | ggctgattca  | ttttttaatt  | 81660 |
| ttttctagag  | acaggctcta  | tgttgcccag  | gctggctca  | aactcctggc  | cccatgcaaa  | 81720 |
| cctcccgtct  | cggcctccca  | atgtgctggg  | attataggag | taagccgctt  | taccagcct   | 81780 |
| ccagttttat  | tgtgttttgt  | tttgttttgt  | ttgagacaga | gtcatgctct  | gtcaccaggc  | 81840 |
| tggagtgcag  | tggcacgac   | tgggtcact   | gcaacctctg | ccttgccgggt | tcaagcgatt  | 81900 |
| ctcctgcctc  | agcctcccga  | gtagctcgga  | ttataggcat | gcaccaccac  | gctgggctaa  | 81960 |
| atttttgtat  | tttttagtaga | gaccgggttt  | caccatgttg | gtgaacacaa  | agtatttgag  | 82020 |
| agctggttta  | atatacagcc  | agctttccag  | aaatgcaact | actgtgttca  | tcaaggga    | 82080 |
| actgaacttc  | gttttcctta  | aaacttatca  | ccagctggtc | atcattttga  | caaattctgt  | 82140 |
| caacaacage  | catgtcattc  | ctgggtatctg | tattggtcac | atctgaacga  | cacacgcct   | 82200 |
| gcatgcagcc  | ctgcaggtag  | tgggtgtata  | acacgcctgg | caagaattcc  | cctccaccct  | 82260 |
| gtgtcccggg  | aacaagaaag  | tgattgtgat  | ccactcacct | cagcctccca  | aagtgtggg   | 82320 |
| attacaggcg  | tgagccactg  | cactcggctt  | cctcccgttt | tttttttttt  | tcaatgctta  | 82380 |
| tattttactc  | taattaactg  | agtcaaaaat  | tgagaatagt | tgaatacact  | ttcatgtaag  | 82440 |
| gcgaatcatt  | tagccgatac  | ttaactctgc  | atgtgggtta | ccatgccgct  | gtgggttagcg | 82500 |
| gggcaagggtg | atgagccctg  | tctcaacaca  | cacaccctgc | ctctcccag   | cccacttaca  | 82560 |
| cgcgtccatt  | cccacgcagg  | tgtgggggct  | ttagaggatt | cctctctctt  | cgtaaagtga  | 82620 |
| gaatgggctg  | gactcggctt  | cactgcccac  | caactccttt | ttttcctttg  | ggaaactgtc  | 82680 |
| cttcccattc  | catataacct  | tcatgggtcg  | agatcactca | ccgagctaca  | ggaagaggcc  | 82740 |
| cattactatg  | gccccatcag  | ggttctgcct  | gggacagtcc | ataaatgctg  | gaagagagag  | 82800 |
| gtgcctttcc  | ctactgaggc  | tgtctaaaagg | agacactgca | aactggggct  | gctgggtggca | 82860 |
| atcttacact  | ctcagtga    | gcctgcctgg  | ctgcagggga | aaccaatgca  | cagaccagca  | 82920 |
| ggggcaacat  | catttgaaac  | cctggagaca  | gctgtgcctg | aagcccacat  | ggctcaggca  | 82980 |
| catgctcaag  | ccactctgat  | ttgcttggtta | cagtcaagag | agtccttggc  | cgggcacagt  | 83040 |
| ggctcatgcc  | tgtaatccca  | gcactttggg  | aggccgaggc | aggcggatca  | cctgagggtca | 83100 |
| ggagttcaag  | accagcctga  | ccaacatggt  | gaaacccctg | ttctactaaa  | aatacaaaaa  | 83160 |
| ttagccgggc  | atggtggcat  | gtccctgtaa  | tcccagttgt | tagggaggct  | gaggtgggag  | 83220 |
| aatcgcttga  | acctgggagg  | tggaggttgc  | agtgaagcaa | gattgcacca  | ctgcactcca  | 83280 |
| ggctagctaa  | caaagcgagg  | ctctgactca  | aaaaaagaga | gtcctgactg  | aagctgagag  | 83340 |
| gctgggtggac | agctgtcagc  | aagcagtgtt  | tgtgagtctg | atgtgagctg  | cgaagtccac  | 83400 |
| acctgatttc  | agagctgggtg | gctctttttc  | caatcaggac | agctccaggc  | ctgagttttg  | 83460 |
| gtgtgggtctg | tacctaacctg | ctgtgtctta  | ggtcaggatc | tctagaagcc  | aagcccaaga  | 83520 |
| caggagttct  | caaataccctg | gatttactga  | gggaatgcct | tccgggggaa  | ctgccagtga  | 83580 |
| ggagagctgg  | aggaggcggg  | caggggcggg  | agctgagcca | ggtgggggtc  | cctgctggag  | 83640 |
| tctagcctca  | gcttgacccc  | acaggggctc  | tggagcagga | acagcacact  | gcattgtccc  | 83700 |
| ttgaggcagt  | cactggctgc  | agttgcctct  | gggagcagag | taaaagtggg  | caggcatttc  | 83760 |
| tgggagtaca  | attccctaga  | gaaggggaca  | gctctgagct | gtgttaccag  | ccaccatttc  | 83820 |
| caggggctgg  | gggatgcgct  | gcactggccc  | agagaggtct | ctgagcaagg  | ccccacaaac  | 83880 |



|            |             |             |             |             |             |       |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| ctccacttca | acctcctgcc  | aaacctcaag  | cagattgaga  | aggagccgct  | agagacacag  | 83940 |
| gaaccctttc | ctcatcagaa  | gtctccaagc  | tttggaggaa  | gaaggaggca  | ggcaggaggc  | 84000 |
| gggaggatcg | tttcaccca   | ggggttcagg  | gctgcagtgg  | actatgatca  | tgccactgca  | 84060 |
| ctactgcctg | gtgacagagt  | gagaccctgt  | ctcaaaaaa   | aacagacaaa  | aaaaatcaat  | 84120 |
| gctcttagca | tctgctgggt  | ccacattggg  | cagtgtgtgt  | gacgtgggtc  | agctgcaagg  | 84180 |
| gaggctggga | aatgcagttt  | cattatctga  | gtactagtgg  | agaaagaaga  | gacgaaagca  | 84240 |
| atgagaatcc | ctgccatgag  | gaagtgagag  | ctgattcatt  | tcagcttcc   | aagggtgagg  | 84300 |
| gatggggccc | agacagggtc  | tgaggccact  | ggaggtcaca  | caggccagca  | gggatacacc  | 84360 |
| cagtggtaga | ccctgggtct  | caggctccca  | ggacacagac  | cacataaaag  | caagggtcag  | 84420 |
| ctcagcacac | agatttaata  | agcactgtat  | gttatgccta  | cagctgcagg  | aaaccagggg  | 84480 |
| gaagatggcc | acagggccaa  | tgccaggcag  | ggcacagggt  | ggggccctc   | aagcagcata  | 84540 |
| tggggaacag | gaaagactct  | gccagcatgg  | tcatgggtgca | cagggtggcta | caatgggtggg | 84600 |
| taatgcccag | ttgggaggat  | gcagtgttag  | acgggggtct  | gtccgatgga  | ttccttaggg  | 84660 |
| ggttatatgc | ttgggaggct  | gatacaattt  | cctagaattc  | tgggattctg  | gatgggtctc  | 84720 |
| gacctgctct | aggggatctg  | ggtctccaga  | gaggtctcag  | ctacaaaagt  | gacctctcc   | 84780 |
| tgaccccacc | ttgtgggaga  | ccccagtgg   | ctcaaggctc  | gtgaataaag  | ggacttacaa  | 84840 |
| ccaagatgcc | tcttgacagg  | gagacggggg  | ctttacagtc  | aggccccagc  | ctacatccct  | 84900 |
| tgccccaggg | gaoggcacac  | aggggtgtcc  | aacctacaaa  | gggcagcgga  | gtgaaagcac  | 84960 |
| cagggccggc | ttggggggta  | agccagcccc  | ttctcaaggg  | gccagtaggc  | cacagaggcg  | 85020 |
| gcccaggcag | gcagggtgga  | gtagctggct  | gtgctatccg  | aggtcgctgt  | cctaatacaga | 85080 |
| gcaggggccg | gagagccagg  | acaggaaagta | tggagagcag  | ggagcgtctc  | tccaggggccc | 85140 |
| tgctgttgga | ggacaccttg  | ggggtgggag  | ccaagtgcag  | acttgaggca  | aaagcctgtc  | 85200 |
| tctgccctct | gcctgcaccc  | ccaccccagg  | cccttcatgc  | ctccaccag   | gcccaccagg  | 85260 |
| gacggaagg  | tgtgaagggt  | acagtggggg  | aggtgccagc  | tcacctccc   | ctccctgcac  | 85320 |
| ctacctgcag | gaggctgtgc  | tccgggtctc  | cctgggggtt  | cagctgggtc  | agcaggactg  | 85380 |
| ggggccaccc | gctggcaaa   | gcctgaatgg  | caccatctat  | ggagggagg   | gtccagggtg  | 85440 |
| gacttagctg | ggaaacccta  | gcacagggtc  | cagcacagg   | gacaaagtga  | gagacccct   | 85500 |
| gctgcaaaga | ctggctcctt  | ggacagaagg  | tgtgtgtga   | ggaataacaa  | aataataagc  | 85560 |
| acagctgtga | cagagtga    | caactgtaca  | ctcttcatt   | ctccctcta   | aaatggctcc  | 85620 |
| cagccaggcg | gcttgggagg  | gggcagtaag  | cgccgcccc   | actccccctc  | cactcccccc  | 85680 |
| cgggcccctc | acccaagcag  | cggttcctgg  | taaagagccc  | ccggaaggct  | tcgcagctct  | 85740 |
| cacgcgggtt | cccgctggct  | ccgcagtcgc  | accaggcgcg  | cacgcgcgcg  | ctcacgttgt  | 85800 |
| ccacgtagtt | aggggtgacg  | gcggtgcctg  | cggggacct   | gagggcgaag  | tcatcgggcc  | 85860 |
| acccggcgcg | agatatcccc  | tggagaaccc  | ccgcccctgc  | ccgatcccc   | gcccgcgcgt  | 85920 |
| cccacgaggc | ccgcgtaggc  | gcgcaggcag  | cgggcgcct   | ggtccagcag  | gcagccgtcg  | 85980 |
| ggggcgctgg | gcgctgggg   | gcacgagacc  | tgaaggccca  | ggaggcgagg  | cctgcgcggc  | 86040 |
| gggaggggcg | tgagccggag  | agaggcccc   | tccccacct   | cgccacggcc  | ccgcgcgcgc  | 86100 |
| ccgcgcgcac | ctgcagaccc  | ggctgcgctc  | cgagaagttt  | aagggctcaa  | ggcaggagg   | 86160 |
| cggcgcgggg | ccgggccccg  | aaaaggcgca  | ggagggcacg  | aaggctctgg  | gcccagcgtc  | 86220 |
| ggcgcacgcg | gggcccgcgc  | acgggcagaa  | gagcagtgcg  | tgggtgagcg  | cgggcggccc  | 86280 |
| gcgggcgaag | aagcggcgca  | gggcccggcg  | gcagcggggc  | cggggacagc  | ccccctgcgc  | 86340 |
| agcccggccc | aggcactgcg  | ccacatactc  | ggagcgcaaa  | cgctggcacc  | gcgcgtccgc  | 86400 |
| cgtgcaggct | tccggccgct  | ccacacatcg  | gttccctccg  | accgagctcg  | ccgacctgg   | 86460 |
| gaaaggcgcg | gggaggctgc  | aggtctcagt  | gcgccccgca  | tgcacacttc  | accagccca   | 86520 |
| tccacgcccc | tggaaagggt  | gggacactga  | aaacccgaga  | aagcatagag  | tgtggcccaa  | 86580 |
| aggcggggcc | gactggcact  | gateggtctc  | ccgtcgtggg  | gatgggcagc  | ggcgggctct  | 86640 |
| ggagggtctc | gcaacacgtc  | agggatggag  | agggagcatg  | gaggctgagc  | caggccttcc  | 86700 |
| ccaaaagctg | aagccccctc  | ttcccaagca  | tccaccat    | tccagaacg   | gccacacacg  | 86760 |
| tccaaagacg | ggtccagggg  | ctcctcctac  | actcaatggc  | actacctgac  | gggggaacaa  | 86820 |
| tgaggctttc | tggggatctg  | gccaatggag  | acttttaaac  | cagagatggc  | atagaagaca  | 86880 |
| gtggctgctc | ctagtctctt  | aatgacaggc  | tgtgacgcc   | tacctttcct  | cctctacccc  | 86940 |
| ctccacctcg | ccgcttatcc  | tccaccatgg  | agaccttttg  | ggtgggtcta  | ttgattttctg | 87000 |
| tttctccctc | cttcttgacc  | tctgtctct   | cctcctgccc  | cctgcagggg  | tgtctccct   | 87060 |
| ctccagccct | tctctgatgt  | cccttgagtc  | gaggtcatcc  | tgccccatct  | gcactggctt  | 87120 |
| cctccagccc | ctccacctt   | catcagccag  | gggagagcgg  | atttatgcaa  | gagctctggc  | 87180 |
| cactgaagaa | cccactgagt  | cctggctgcc  | aggggttgga  | ttcagagaag  | tccctgacga  | 87240 |
| cacaccgccc | ggctttgtcc  | tgagccacag  | gacctaaagg  | ttcatcagtg  | gttttagtga  | 87300 |
| ccaggcccca | tgatagggtga | cacctgtctg  | gagacagaga  | aggtgtctgg  | caccccgatc  | 87360 |

32/154

|             |             |             |            |             |             |       |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| acacaccagg  | ggttgggtag  | acagaggggtg | agctctgccc | acatgcccac  | aggctgtgtc  | 87420 |
| taaataagagc | caggataaact | ggaagttggg  | acttggcagc | cccttcccag  | gatttgagtc  | 87480 |
| acaaaaacca  | ggctaatacga | atcccccca   | gcctccacaa | gttctccagt  | ctggagcatc  | 87540 |
| tcattcagac  | tagaagcctt  | tctgggtgcc  | agcctgagtg | gacgtgcaca  | gcctggcccc  | 87600 |
| tcattccatct | ccctgtgtgt  | ttgtgcacag  | gggctacggg | gtggtatccc  | ccccagggc   | 87660 |
| tgggagcagg  | gtgcctctct  | ggcaatggct  | gtgtctttca | ggctgtggat  | tggaccgcgt  | 87720 |
| ggtgagagt   | attcctgctg  | tgacttcagc  | acctgcctct | gcaggaagga  | tgaacatagc  | 87780 |
| tcgagggctt  | agaatgggtcc | ctgagtggag  | gggatagatg | gcagcggggg  | gagtcctccc  | 87840 |
| ttcctccctt  | ccccaaagg   | cacacttttg  | tttctctca  | cagccttgcc  | aggcacatcc  | 87900 |
| tcaggtctca  | gctgcacaca  | ccagcacctg  | atgttgccca | agggtttgaa  | cccacctctg  | 87960 |
| cctgttgctt  | acctgtcacc  | cttgcccaga  | acaccttac  | ccagtctcca  | ccctcacagc  | 88020 |
| agctcatctc  | agctgcaaga  | tctcctccac  | cttaaaccct | taccagattc  | cactttccac  | 88080 |
| actacactcg  | gtaaagtga   | ccctccctcc  | gcgacatcca | tgcagggacg  | gagcagtcga  | 88140 |
| agaaggtgga  | acaacagatg  | aacaacaccg  | aagattcacc | aaggaatctg  | tgtgatgggt  | 88200 |
| aggagggctg  | gggctgaggg  | gagggtgaga  | gagttgggag | atggcagaca  | ataggcaaaa  | 88260 |
| gttctctggg  | cacgggtgcc  | aaggctggac  | tcgggagggg | gatggcgggt  | gtggcttgac  | 88320 |
| ggggctggga  | ggtgaggtgc  | tcagcacata  | gtggggctca | ttgccagggg  | ttcccaagcc  | 88380 |
| gtgcgcacag  | gcctctcaact | ggcaaccctt  | cctgtgggtc | agcttttcc   | ggaatatctt  | 88440 |
| gccctcccac  | cttgtcactg  | aggagcaggg  | acagagaggc | cagggttcgg  | ggtccagggg  | 88500 |
| aagagagggg  | cgctcaccga  | gtaacagcag  | cagcagcagc | gcaggcccca  | ggcagcggac  | 88560 |
| catctaggac  | cttcaacaga  | gaaggatggc  | tggcagagtc | ctagtctgat  | aggcgcccc   | 88620 |
| ttcctagtgc  | ctctggagca  | ggagcacagt  | gaaaaccag  | ctgtctgcgc  | tgataccctt  | 88680 |
| gggaggagcc  | tttgctgcag  | caactccccg  | ccctcattaa | gtctccaccc  | caggctgcac  | 88740 |
| cgtcaggtaa  | acctgggagc  | cacgggatcc  | ggcgctgaa  | tgtgcctaga  | cggaccccaa  | 88800 |
| cacactagcc  | ctgccacgca  | caccagcaa   | cacacagaca | accacccagg  | gagacctggt  | 88860 |
| ctctcccatg  | ctaaactagg  | aacgctatga  | aacagcattt | ttcttttcc   | tttttatata  | 88920 |
| attctatttt  | aaacttaaaa  | aaattttttt  | gaatagagtt | ggggtctcgc  | tctattgcgc  | 88980 |
| aggctggtct  | tgaactcctg  | ggctcaaacc  | atcctcccac | ctcggcctct  | caaagtgcta  | 89040 |
| ggatgacagg  | tgtgagccac  | catgcctggc  | ctatttttaa | cttcttattg  | ttgactaatt  | 89100 |
| ttagacttgc  | agaaaagtgt  | gaaaatgtaa  | gcattttcct | tacacccttc  | accagcttc   | 89160 |
| ccctgatggt  | ccatcttact  | taatcatagt  | ccaatcatta | cagctaggaa  | actaaccttg  | 89220 |
| gtacaatcct  | gttaacgaaa  | ctgtagactg  | ttttgcattt | ctcatttttc  | cactaatgtc  | 89280 |
| ctttttctgt  | tccaagatcc  | attctaggat  | cccacatcac | agttgtctcc  | ttattctcct  | 89340 |
| caatctggga  | gttccctcat  | ctttccttat  | ctttaccctt | gacacttttg  | aagaatcctg  | 89400 |
| gccagttatt  | ttgcagaatg  | tttctcttga  | gttgtctgat | gttttattct  | gatcagaatg  | 89460 |
| agacacagca  | ttgttttgac  | taacccaaaa  | gttattctat | aagtaaatat  | tgagggtaaa  | 89520 |
| aatctcatcc  | aacctgggca  | acagagtaag  | gcctgtctct | aaataaagtc  | tcaacactaa  | 89580 |
| gatttaaaaa  | gtgaccagaa  | aagccccac   | tatgatttgt | cttgactttt  | ttttaaaaaa  | 89640 |
| caaacacaaa  | aacaaacaaa  | caaacacaaa  | cgatgtcttg | ctctctctct  | ccctcaggct  | 89700 |
| ggagtgcagt  | gccaacgatc  | tggctcactg  | caacctccgc | ctcccagggt  | caagcgcttc  | 89760 |
| ttctgcctca  | gcctcccagg  | tagctgggat  | tacagacgcc | cgccaccatg  | cccggtctaat | 89820 |
| ttttgtattt  | ttagtagaga  | caaggtttca  | ccatgttggc | caggctgggt  | tcgaacttct  | 89880 |
| aacctcaagt  | gacccgcccc  | cctctgcctc  | ccaaagtgtc | ggaattacag  | acgtgagcaa  | 89940 |
| ctgcgcttgg  | cctcattagc  | atcttaaatc  | tcacacacag | ggtgtgttcc  | ttactgttat  | 90000 |
| aaggagcaaa  | ggatcagttt  | gaggacaggt  | aaaataaaaa | tgcgcttgct  | gcctagaggg  | 90060 |
| agaagtccct  | gctgaagata  | gctttgcttg  | aatgagctca | attgcaatgc  | cagtgtctgag | 90120 |
| gcttggtgac  | tgtacgggtca | ccacagttgc  | tgctgcgcgc | ctagaacatg  | gtcactttct  | 90180 |
| tgactacctc  | tccgtgtctc  | gtacatctgt  | ctgtgggttg | tgggtggtcca | tttccctaatt | 90240 |
| tttttaatga  | atcagaagac  | tgtgatgtgc  | tttccgctgt | gctaaccatg  | gcccgtgaag  | 90300 |
| caaaatgtaa  | accaagatgc  | ccctgcagtg  | gttgtgtctc | actctacgac  | atctgttacc  | 90360 |
| ggaaaggggt  | ccagattcag  | accccaggag  | agggttcttg | gatctcgtgc  | aagaaagaat  | 90420 |
| ttgagacgag  | tcataaaagt  | gaaagcacat  | ttattaagaa | agtaaaagaa  | taaaagaatg  | 90480 |
| gctactccat  | agagagcgca  | gccctgaggg  | ctgctgggtg | cccattttta  | tgggtgtttc  | 90540 |
| tggatgatct  | gctaaacagg  | ggtggattgt  | tcatgtctcc | ccttttttaga | ccatatatgg  | 90600 |
| taacttcctg  | atactgccat  | ggcatctgta  | acctgtcatg | gtgctgggtg  | gagtgtagca  | 90660 |
| gtggggaccg  | accagaggtc  | actctcatca  | ccatcttggt | ttgggtgggt  | tttagccagc  | 90720 |
| ttctatattg  | caagctgatt  | tttttggttg  | gtttgggttt | tgagacggag  | tctcgtctca  | 90780 |
| ggctggagtg  | cagtgcacatg | atctcagctc  | actgcaacct | ccacctcctg  | ggttcaagcg  | 90840 |

|             |             |             |            |             |             |       |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| attctctctgc | ctcagcctcc  | caagtagctg  | ggattacagg | cacacaccac  | catgcctggc  | 90900 |
| tagtttttgt  | atttttagta  | gagatgaggt  | ttcaccttgt | tggccaggct  | ggtctcgaac  | 90960 |
| ttctgacctc  | agggtgatccg | cccacctcag  | cctcccaaag | tgctgggatt  | acaggcgtga  | 91020 |
| gctaccgcgc  | ctggcctact  | gcaaactatt  | ttatcagcaa | ggtctttatg  | acctgtatct  | 91080 |
| cctatctcat  | cctgtgacgc  | agaatgctgt  | aactgtctgg | aaacgcagcc  | cagtaggtct  | 91140 |
| cagccttatt  | ttgctcagcc  | cctattcagg  | atggagttgc | tctggttcac  | aggcctctga  | 91200 |
| cacatctctt  | tgtgttttgg  | cgtgtgggag  | gaaagagggg | tgagggaagg  | aaactcaaaa  | 91260 |
| ccaagctctg  | accacacagg  | gcagggtacac | tctccacact | gtctgtgggt  | gccacaagtc  | 91320 |
| aagggagggg  | cagagagaga  | agaagggtgtg | acagatggcc | gcaggccaca  | gaatgtcaga  | 91380 |
| ggaagcccag  | ttcctcccgg  | ggcagcccaa  | gtagctggta | ggtgggtggc  | caaacagagg  | 91440 |
| gcgtcacagc  | tgagctgggc  | tcgctcgcta  | ccccagctc  | agcgtccact  | ctgcccctca  | 91500 |
| gtacctcttg  | ctcagcctca  | gggtccatgc  | ctaccctcct | gcttcccagt  | cacttctgcg  | 91560 |
| tgccctctgc  | ttttctgctg  | ttggcccat   | gccagctcct | ttctgctgag  | cttctcttct  | 91620 |
| ccagttctgc  | agcacagcca  | ggtgatcctg  | ggctccagac | aggcctctcc  | cccagttctg  | 91680 |
| ggcctccctt  | cttagagccc  | tctttccttc  | ccacgtggcc | tccccagggt  | tcgccactga  | 91740 |
| atggagaagg  | ggtggagggg  | gtgctgggca  | gtcttggggg | ttagccaaga  | gggcagagtt  | 91800 |
| ggcctcccca  | gggtcccttg  | tagctggagt  | cccgcggggt | ctagacaccc  | cctcctgaag  | 91860 |
| ggtaagagcg  | ggggaggtat  | attaacgtgt  | atttttagag | tctctctttt  | tttttttttt  | 91920 |
| ttttttgaga  | cggagtcttg  | ctcttgttgc  | ccaggctgga | gtgcagtggc  | acgatctcag  | 91980 |
| ctcactgcaa  | cctccacctc  | ccatgttcaa  | acaattctcc | tgccgcagcc  | tcccagtag   | 92040 |
| ctgggtattat | aggtcacgtgc | caccacaccc  | ggctaatttt | tgtattttta  | gtagagatgg  | 92100 |
| ggttttacca  | cattggccag  | gctggtctcg  | aacttctgac | ctcaaattgat | cctccgcct   | 92160 |
| tgccctccca  | aagtgtctgg  | attacaggcg  | tgagccacca | tgccggcct   | tagattctct  | 92220 |
| attttatgat  | ggtttaacat  | ctcggggtgg  | gggcttgttg | gctggagaga  | aactgcttga  | 92280 |
| ttcctggaga  | tcagaaacaa  | ctcatgcctt  | tcatatgcaa | accgaccagt  | cttgagtcca  | 92340 |
| tacaccaacc  | accccttcca  | aggaaactct  | acatacgaaa | ccagtatttc  | ccctgcccta  | 92400 |
| aaccagctca  | gggccaggca  | cctaccaga   | caattagagc | ccaacccccc  | gccccaaaac  | 92460 |
| ccaccagaat  | gattcaaaaat | gccaatccta  | ccctcttccc | ctggcctgcc  | ttgcctcccc  | 92520 |
| agtggaaaacg | gcactgtggg  | ctgtggcctg  | tgccctccac | tcgctcctga  | ttctgtccct  | 92580 |
| ggaccaaac   | tagtgctccc  | ccactgtggc  | cctgcattgt | gggaaactgt  | gagtaataac  | 92640 |
| ttatttcaac  | agcatgggcc  | tctgtgtcat  | cagtcacctt | cataaattaa  | aattttgcaa  | 92700 |
| ctacaactga  | ggcaggaggg  | ggcttttagag | agcactctca | cctggccccc  | tgggcttgtg  | 92760 |
| gagtggagcc  | cgagggtgaag | ttgcctccct  | gcactcagtt | ttgggatggg  | tttgttatgc  | 92820 |
| ggcaacagct  | ggctgatata  | ggccactcag  | cagctcttgc | atgaagcaat  | gggagaatgt  | 92880 |
| gaacgcccac  | gggaggcagg  | agtgcacagag | caaagagggg | gttcaaaacta | ggcaacaccg  | 92940 |
| tcctgtgccc  | agacacatgc  | ctggggctct  | ggctgccatc | taggctcggc  | ctgccagcca  | 93000 |
| ccaccccgct  | agccaccacc  | ccaccaacca  | ccaccccgcc | aaccaccacc  | ccgtgtgggc  | 93060 |
| ctcagggcac  | cccatgggtt  | gaatttacat  | acagaaaaga | ctaaccaagg  | tccaagtgtg  | 93120 |
| atatgtcttt  | ttaaaattta  | tttttagagac | agggtcttgc | tctgtcacc   | aggctggagt  | 93180 |
| gcagtgggtg  | ccaatagct   | cactgcagtg  | ttgaactcag | gctcaagtga  | tcctctgct   | 93240 |
| tcagcctcct  | gagcagctag  | gcctacaggt  | acacaccacc | acgccagct   | aatttttaaaa | 93300 |
| tttttctgta  | gatgcgggtg  | cttgctgtgt  | tgcccaggct | ggtgtggacc  | tcctggcctc  | 93360 |
| tggtgatcct  | cctaccccag  | cttcccaaag  | tgctgggatt | ataggcatga  | gccaccacgc  | 93420 |
| ctatagccaa  | catgtctttc  | ttttgacttc  | tactttggta | tcttttctta  | aatgggtccc  | 93480 |
| tctgtccccc  | cgacacacac  | agaatggggg  | agaggctgtc | agattctgag  | ctccagaacc  | 93540 |
| tcagggtgtag | cactgggatt  | gggggtgggg  | gctcaggaac | cacctagggg  | agaagacagg  | 93600 |
| gtgggaagaa  | acaggaagga  | aggtccccaa  | aattatgttt | gtttgcagag  | gccagccagg  | 93660 |
| ctccagggga  | gtgtggactc  | agtcgaacca  | tagggcccca | ggaccactag  | cttctggcca  | 93720 |
| gcagtcattgc | cctccacaga  | gctgggtccg  | tggaaattgc | atgtaggaga  | cacaccagac  | 93780 |
| tcacaggaca  | gagccctttt  | gggatggcca  | gcactaccca | gcctccactg  | gtgagggagg  | 93840 |
| tcaggggctg  | tgtgaccttt  | gcttctggga  | ctgatggttt | attgagctgg  | agagtgtgcc  | 93900 |
| cagcagtggt  | ctccagccct  | caggaaacttc | tagtgtggct | ctgggttcc   | ggagtgggtg  | 93960 |
| ggtcgaagct  | ccactcgggg  | aagaaacttc  | caagctgcct | gcagggtgctg | gaggtccggg  | 94020 |
| gattcactgg  | ctctgcccc   | gcagttcaag  | ttcctggagt | ggctgtcagt  | ggccacctgt  | 94080 |
| ctttaaatct  | gttcattttta | ggagctacct  | ctcaccagag | gcaggatctt  | ggcatctgga  | 94140 |
| cttgatctgc  | tgagaatgag  | gaggatatgt  | tgtcccctaa | ggactggggc  | cccaggctgc  | 94200 |
| aagctgtgtg  | gcagagagcc  | catcctcact  | cagtgaggac | cagtgatcca  | ggaaaagcca  | 94260 |
| cagcttctcc  | ctccccagcc  | caggggcttc  | cagcatcctg | gtctccatga  | taaccaagag  | 94320 |

34/154

|             |             |            |             |             |             |       |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| gtcataaact  | catttccata  | ataacctgag | cccagaaacc  | tgattagggg  | gcagcaaact  | 94380 |
| gaggggtggg  | agaggtggga  | gggtgggcga | tgagaggggg  | aggctttgaa  | tccaggtccc  | 94440 |
| tgcttacctt  | gggggtcagg  | cgagactgct | ggcagaggct  | tctcaggggtg | gctgctgggc  | 94500 |
| tcatgagagt  | tctcaggggtc | tgggagaaat | gggtggagggt | aaatggtgtg  | aatatggtca  | 94560 |
| gcaggagacc  | ctggggctgg  | ggaggggcat | aggggactca  | agggtgactgg | gtgetgccc   | 94620 |
| tctggaagga  | ggcaggaggc  | atgagccctt | cccttctccc  | ttccctctcc  | acctccccct  | 94680 |
| ggtgcctcac  | tcacccaggg  | gccagggctg | tccagtggct  | gtggggcccc  | actccatggg  | 94740 |
| gtgaacgcgc  | cccaggggggt | ggtccctgtg | tgggccatct  | ttggggctga  | gcaacgtgat  | 94800 |
| aagagtccag  | gaggttggca  | cagtgatect | gagtgggtta  | ttgcctcccc  | gcagcatggt  | 94860 |
| gtccagcccc  | gggagttctg  | cgtttactga | gtttcttggg  | gcacccatct  | gctccaagtc  | 94920 |
| acctctcacg  | ctcccttctt  | gctccctctt | caggggagcc  | ttgggatcca  | ggctccaagt  | 94980 |
| gagcctcatg  | ccctcggtcg  | gcacctcctc | tctctagtcc  | taacatttcc  | tccaggtctt  | 95040 |
| gacaccaccc  | agcagcctgg  | cactctccag | atgctggcat  | cgctcagctt  | ccaaagaacc  | 95100 |
| ttggatgtcc  | gccccctcgg  | cagctatgtc | tgctctcctt  | gccccctggg  | gccccctgtc  | 95160 |
| ccttgatgat  | tccaagccat  | ctttgactgt | ccccatcccc  | tcccccaagg  | ccttgatcat  | 95220 |
| tctgtgatg   | ttccttcaaa  | acatttctcc | ctgcccctgag | actcccgcct  | ggggatgaga  | 95280 |
| agcagccgcc  | acctcttgca  | gcgccccctc | cgtgctgaca  | cgccaggctc  | tggccacctt  | 95340 |
| gctcctctgc  | ccacagaccc  | tcaatcacaa | ctcgctttgt  | cagggtcctc  | ttagctgcca  | 95400 |
| ccggggggcc  | agggtggtgcc | ctgccccctg | ctggtatgcc  | ctctgcccc   | atctctggcc  | 95460 |
| caaatcatgc  | catctccctt  | ggcttgcccc | gagcactccc  | aagaccaggc  | tatgtcagac  | 95520 |
| atggccacag  | agtgcttccc  | ctgcctaggg | cctgggtgca  | gggtgagtcc  | taggcagacc  | 95580 |
| atgcttagta  | ttatgtgact  | ccccactccg | ccaccaccca  | ggtcacagag  | aactgggtta  | 95640 |
| aggcagggcc  | ctggcacagg  | ggcagccagc | accgcagctg  | accagtggta  | tggagtga    | 95700 |
| agatgtgctg  | ggcccagcat  | ttgggaactt | caaggggggtg | acagaggtga  | tttgtgcaga  | 95760 |
| ggaagtggca  | aagggccgaa  | aactggtgag | acagaggctg  | gacaggcctc  | cgggggcagc  | 95820 |
| atggtacagg  | gactgcaatc  | tgagccaggg | aaaaacaggg  | cgaagtcaag  | ggtgaggcag  | 95880 |
| ccagctgggtg | ggagaagcag  | gagagtggac | aagaggagct  | gtactgggag  | gtagagggcc  | 95940 |
| atgccttgcc  | gtgctgggtg  | tggggcaggg | atccaccccc  | tcgcttgact  | gggaggccac  | 96000 |
| tggaaacctc  | tgttcaaaag  | tacttctttc | catggcctct  | ggggctgctc  | tctgcacctg  | 96060 |
| ggggcgaagg  | tcaggggcctg | ccccagctcc | cacagcccca  | gcagagctct  | gaggagggga  | 96120 |
| accgcaggag  | taggctcagg  | aagcaggcgc | tcggagccta  | cccactgcac  | gcagggtccc  | 96180 |
| ttctgcagcc  | ccagctgcat  | cgctgcagat | gggctcctgg  | gagtcggtag  | caacaccagg  | 96240 |
| ccaggccggc  | ccctgggagc  | agaggcagca | ggacgctgag  | gagcatggcc  | agcagggaagg | 96300 |
| tgtcatggte  | tgcggggatt  | gggggaaggg | gcgctgagtc  | ctgagcagggt | gcaccacccc  | 96360 |
| agctcctgcc  | cacatgcccc  | ccactggcat | actttcagcc  | tgcacagggc  | cactgtccat  | 96420 |
| gctgccacca  | aagcctggct  | tgtcacagaa | gggtggagcc  | cagcctggag  | cacagtggca  | 96480 |
| gttatggttg  | ctattgcaaa  | cctgcagaga | agagaagagg  | agggtcacgt  | aggattagga  | 96540 |
| accccaagg   | caccccact   | cctcgggctc | tcaccccggt  | gctgtggcag  | gcagttaggc  | 96600 |
| agcgctgaag  | ctcctggaag  | gcattcttcc | tcgacgcctc  | gctctggcac  | acctcagggc  | 96660 |
| agtgaccagg  | gcagtggagg  | ggacactggc | tgcgggatt   | caaacggcaa  | ggaggggctg  | 96720 |
| gggtgggcaga | gctcaccatt  | ctaggtccac | actgggtgcc  | tggctctacc  | aggcccaggc  | 96780 |
| caagcaggtc  | cagctgggca  | ctggggagtg | ccaaggctcc  | ccgacaagtc  | acttctctggc | 96840 |
| catctaggtg  | aacggtagag  | tccactggca | ccatgtgcgg  | tgcgagcagg  | ctgggctttc  | 96900 |
| caccctggca  | ctgcagcttc  | ccacacaggg | catccctggg  | gaggaagtag  | aggggggtca  | 96960 |
| acagctgcag  | tacccccctt  | ccccaaaccc | actccatagc  | ttctgctccc  | tccactcagc  | 97020 |
| tccactccct  | acctccctgc  | acagggcagg | aagtggccct  | cgctgtcctg  | gcccaggttt  | 97080 |
| ccatgagcat  | ctcccgca    | gttcaccacc | tggaaacagg  | cctcgggagc  | tgggtgggag  | 97140 |
| cctgaggaag  | catggggcag  | gctgggggca | gctcggagag  | gggctgcgct  | cagcgggcct  | 97200 |
| cagtttcccc  | tgcccatctt  | ccccacagta | gaaaactggc  | cccaccagag  | gatggggggc  | 97260 |
| aggggtgcaag | ggtgctcgtg  | tcctctcacc | aggccccag   | agctgctggc  | actgctgctc  | 97320 |
| cagcgtggga  | catgcgccat  | cccagcagta | gccactgccc  | ctggcacagg  | gtgagccgtc  | 97380 |
| cagtaggtaa  | acgtctgggg  | gacagtggga | ggaggtgccc  | gtgcaaaact  | cagggaggtc  | 97440 |
| acagtcaccc  | atggcctggc  | ggcacagcgc | tccagccggc  | ttcagctgcg  | caggtgacgg  | 97500 |
| gtggtgggga  | aggcagagag  | aggccacgtg | cagtgaagag  | tccatgccga  | gagcgcggt   | 97560 |
| cggagctggg  | ggagccaggc  | ctacccaagc | ccagcaccca  | acgggggaac  | ctgagggcac  | 97620 |
| caattaacta  | aggccaacag  | gccggctccc | aagctccccg  | aaaccctcac  | cctgaacctt  | 97680 |
| ccatgccttc  | accaggcagc  | gcacgcagca | gtccccgtgg  | gcgcactggg  | ccccggggcg  | 97740 |
| cagcgagcag  | ttgtgagcaa  | agcagcagag | gtcgcggcac  | tcctgggacc  | agaaaggcaa  | 97800 |

35/154

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gaagggccca  | ggtgagggcg  | cagcgcccca  | gacctgagcg  | gagagggcaa  | gtggggggccg | 97860  |
| ggcgagccga  | cttaacctgg  | ccagggccgc  | agtcacactc  | ctcgcccgct  | tccacgaagc  | 97920  |
| cgttcccgc   | gagcgccggc  | ggcaccggga  | gtccggggtc  | cggggcattg  | gagaggcaag  | 97980  |
| cgcgcgcgc   | cttgcggaag  | aaggcgcgca  | gctggcgggc  | gctgcaggcg  | ctgaacacgc  | 98040  |
| gcggaacgg   | gtgcctaccg  | gcacggggag  | ggcattgggc  | atggagggac  | agtcccccaa  | 98100  |
| cccccgcgct  | tctctgatcc  | ccacccttgg  | gcttggttac  | agccgccaga  | cgcgcagagc  | 98160  |
| ccagagaggg  | gaagtaacct  | gcgcaaagtc  | acacaacaag  | cgggacaggg  | gacgatgcgg  | 98220  |
| ccccaatagt  | gagcagcccg  | ggaccaagg   | tggaatcgcg  | acccgacggg  | gctcctcccg  | 98280  |
| gtgtaggagt  | aacctcgcca  | ggttactcgg  | aaaataatct  | tcataccggt  | gagaatccac  | 98340  |
| tttgccctgag | cttcttccct  | ttaagcctca  | taaaccaccc  | tgaagcggac  | actatgatca  | 98400  |
| ttatccccc   | tttacgaag   | aggaaactga  | gggacgacca  | aagaaacgca  | gcggagggaag | 98460  |
| tccccaggac  | tagccgcccc  | gccgcagccc  | cgacccccca  | cccgcgtacc  | cgggtggccgc | 98520  |
| agccatgacg  | cagcctccgg  | actcgccgcg  | agcctccacg  | cagcagccgt  | cggggtcgtg  | 98580  |
| gctgagggcg  | aggctgtggc  | cgatctcatg  | ggccatgggtg | gctgcggcg   | cgatggggag  | 98640  |
| ctccgagtg   | tccctggggg  | ccgtggggag  | gcggctactg  | cggccgtaga  | gcctcctgtc  | 98700  |
| tctccctcgc  | ccccgcccgc  | ggggctcacc  | gtgctcacgc  | ctcccgagct  | ctcggcgcg   | 98760  |
| cacatgccct  | cgacggggcg  | caggccccact | gtggcgccct  | ggaaggcgcg  | gcccctgggg  | 98820  |
| gcggagcgcg  | gcgtgaccag  | gcggggcccg  | gaggtgaggc  | cgccccaccc  | gggacccgcg  | 98880  |
| tccgggtcag  | aggcaccac   | gtgagcagct  | gcgcggagtc  | gtggggccgc  | tgcgccaca   | 98940  |
| gccccggcg   | ccactgcagg  | aaggcccaga  | gcgtggcggt  | ggcgctcctg  | gtgacgcggc  | 99000  |
| tgcggtcccg  | ctcgggtccac | acctccaggc  | cggtcagcgc  | cacctgaatg  | tccagagctc  | 99060  |
| tgagaagctg  | agggcgaggc  | ggggctgaag  | ccgggacagg  | gcgccccatc  | gcgcccgttg  | 99120  |
| tccttcgtgg  | ggcgcccttc  | ctcttcccca  | aacccaccca  | gcacctgctt  | gtcctgccgc  | 99180  |
| cgccaccccc  | atcacgcgtc  | tctccccgcc  | gcccccaacc  | tgggtccacgt | agttggcgac  | 99240  |
| ttccaggaga  | cgctgtttgg  | tgtggttcaa  | gtttcggtgc  | cgagtcaaga  | actgggaagg  | 99300  |
| cagaaatccc  | ggtggcttga  | ggggctgagc  | tggccccatc  | cctgaccccc  | ccaacccctg  | 99360  |
| gggtctctcc  | tcaccagggt  | gtggtctgcc  | acaatgtaca  | gttccaggta  | cttccgggtc  | 99420  |
| ctgcgcgctt  | ctcgccctgc  | ctgcggagg   | gcaaattggg  | accctgagtg  | gaagctgctg  | 99480  |
| ggcttgagcc  | ctgaccccca  | acccagctc   | ccagaaggaa  | gtttaacatg  | ttttctggaa  | 99540  |
| cttgtttctt  | cagacttcaa  | taaaaatact  | gggactcgag  | gcctgtgaat  | tactgtctct  | 99600  |
| tctgatttgg  | agggctatag  | atacagcatt  | cccactccca  | tccgatcgat  | gcccctgacc  | 99660  |
| ctgctctggg  | gaccaccagg  | aaggctggtc  | atgcccgctt  | tgttcccagg  | atccctgtgg  | 99720  |
| ccacaggttc  | ctttccagg   | gagcagctgc  | tccatccgaa  | agatctcgtg  | ggttgagaag  | 99780  |
| tccttgagag  | cccggggtgg  | ccagggacgc  | agataatagc  | tggcattcct  | gctgaggggtg | 99840  |
| atcaggccac  | taggggtgcag | aggggttagga | gcgggtgtga  | gggagctctt  | tccccatccc  | 99900  |
| aggcccagcc  | tccctctcca  | gagctcacct  | catcccagag  | cagggtgcaga | ggactaccca  | 99960  |
| ggagtccggg  | aagcccctta  | ctcgcccttg  | gtagtggcaa  | tgatcctagg  | gaggaagggg  | 100020 |
| ccagcccca   | atctcagcca  | gggctggagc  | aagaggggca  | agagggagg   | tgtggtagg   | 100080 |
| gctggtccca  | acgcgccctt  | aggaatgcaa  | ggaggagttag | gggtagggaat | ggtggggggg  | 100140 |
| tacctctggc  | ggtgcattccc | agagcccatg  | gaagcatctc  | accgtgtggt  | tggggggccag | 100200 |
| caccactggc  | tgcccatctg  | ggccgtagtg  | ggtttctatg  | tatcctgggg  | ccagcagcct  | 100260 |
| gctgagaggg  | ggtgttacag  | ggaacactga  | attcagcttc  | ctcctgcctc  | ctccaggatg  | 100320 |
| tctcccagcc  | ttcctcccta  | aatgctaattg | gagcagcttt  | atgagtgaga  | cactcacagt  | 100380 |
| gtgtcttagg  | gaagggacag  | gagcaatgg   | gacttgctca  | gatcagaaac  | tcttggggct  | 100440 |
| agaggaagga  | gccttggtga  | tggcttagtt  | gtgggagatg  | tgaatatggg  | aagcaccagg  | 100500 |
| gaggacgccg  | gggaggagtg  | ggaatagggg  | aagagtttgt  | ggttcccagg  | ggacctgcag  | 100560 |
| caggcagcag  | gatccacagg  | atcgggagg   | gaggagtcag  | gagacactgc  | cgaagaatgg  | 100620 |
| gacttggagt  | tggggaaatg  | cggtgacctc  | ccccagttcc  | cctgcctgct  | gcccctcttt  | 100680 |
| gttgggcata  | tgggtcgaccc | tcttgccccc  | acctgcctta  | gatccttgaa  | atattttctc  | 100740 |
| cagacttcta  | gacccacat   | acctccacc   | tgtccttcag  | tgattgatgc  | tcaccccttg  | 100800 |
| cctccagaga  | aaacagaatc  | gccacctgcc  | cacgtgctt   | ccacctccc   | tgccttctcc  | 100860 |
| acacccactc  | cggtaatgat  | tccatcttca  | ggctccatct  | caacaggatc  | tttcccacac  | 100920 |
| ggatggatca  | atcataagtc  | aatctgtctt  | ctttaagaa   | aatccttaac  | ccaacctcac  | 100980 |
| cttggcctca  | ttacctccag  | accaccgcgt  | aatgatggct  | gcttcccccc  | tcccaggcat  | 101040 |
| tcaccacct   | gccccagctc  | tgccccctac  | ccctgcccc   | cacacacccg  | ccacctagg   | 101100 |
| aggtaggtga  | tgtgaccacc  | ccattgaaag  | ggtagggacg  | tcgggaaaat  | atggttgggc  | 101160 |
| acagtggaac  | tagagtttgt  | tccctgtcca  | tccgactcca  | cgaggagaa   | taaaatacgt  | 101220 |
| gtcaagtgct  | cagaacagcg  | cctgcggtca  | agcactcagt  | aggtgatata  | tactgataac  | 101280 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ataatctggg  | tgggttttaag | agcctgcgct  | ccagcccggga | cacccacccc  | accagccca   | 101340 |
| aggcaccttc  | ctgagcacag  | gtctgcctgt  | cccttcccca  | gtcctaaaatc | tttggctctg  | 101400 |
| gatgggtccag | gacactgagc  | accaaaatgg  | ccttctgtga  | tctggccctc  | ctgggactct  | 101460 |
| ccaaattcat  | tccccctgc   | tccccctcct  | gtagatggag  | accttccagg  | caggtcagac  | 101520 |
| aaactgctgc  | ccccaaagtgt | agccactgca  | tctttttctt  | ttttcttttc  | tttctttttt  | 101580 |
| tttttctttt  | tctttttttt  | ctcttttttt  | gagacggagt  | ctcactcttt  | gccagggccg  | 101640 |
| gagtgggtgca | gtgggtgtgat | cttggctcac  | cacaacctct  | acctccgggg  | ttcaagcgat  | 101700 |
| tctcccgctct | cagcccccta  | agtagctggg  | attacaggcg  | ccgccaccac  | gcctggctaa  | 101760 |
| ttttttgtat  | tttttagtaga | gacgggggtt  | caccatgttg  | gcaaggctgg  | tctcaaactc  | 101820 |
| ctgacctcag  | gtgagccgcc  | cgctcggcc   | tcccaaagcc  | accgcactct  | ggtccctgcc  | 101880 |
| attcccttag  | cctgggggtgc | cggctcatct  | tttccctcta  | ggatttcttt  | agactcagca  | 101940 |
| tatcttgcaa  | atgtccacta  | ggtgggtgctc | actcatcgcc  | agcaggggagc | taacaagccg  | 102000 |
| ctcctgggggt | tgggaggggcg | gaggtgcccc  | acagcggggc  | tgacagcctc  | agcggctctc  | 102060 |
| ttcagcctcc  | agggagccaa  | ccacaggcct  | gcgtgactct  | ccctgtcatc  | tgcacctctc  | 102120 |
| ctgggggtcct | ctgcccattcc | agccaccgc   | acagatctgt  | gtcagtcctc  | gcccccaac   | 102180 |
| actgatcccc  | tctcccagc   | cctaccccag  | cctggcactc  | actggttctt  | ctccagctca  | 102240 |
| agcaggagct  | cctggccttc  | agcctccagg  | gccaccagcc  | ccatgtcttg  | cttcgagacc  | 102300 |
| tgggcaagaa  | agagtgtgga  | gctgagatgg  | tggcctccag  | gcctcctgcc  | tgccaggggag | 102360 |
| taggtggcct  | gtggagccgg  | ctggggagga  | agttcttggg  | gagaaagtgg  | gctggggagt  | 102420 |
| cagcaggacc  | ccccacatac  | tatggagggc  | gtggaggagg  | tgagaacata  | caaagatgtt  | 102480 |
| cccaaactca  | ggatgtttgc  | agtcctgaca  | acagccactt  | ggaaggggcgt | tggcacagcc  | 102540 |
| tgcagggcac  | accagcatcc  | tccctagaga  | ccagggtcc   | cagaaagggtg | cccctccct   | 102600 |
| ggccccgccc  | tcttctttca  | tgcccagaag  | gggcatcaaa  | agcaggggaa  | gacagagggg  | 102660 |
| tgtctaggac  | attatggggg  | catcgggtag  | ccatggctag  | ggcctcctca  | gagcctctgc  | 102720 |
| tacctgaggc  | ttggttccaa  | atgagctgct  | gctcatttcc  | tatagaattc  | aaatttgact  | 102780 |
| cctccacttc  | caattttggc  | aaactgctcc  | ctcttccaaa  | gttttcctgg  | gcctccagca  | 102840 |
| gcccccgccc  | ctccggctcc  | gacacctgct  | tactggacc   | cacgaagtaa  | acatggacgc  | 102900 |
| cattccagcc  | aagagagcac  | actggctctc  | agctagggtg  | caggaggctg  | gcttggacgg  | 102960 |
| ccagccctct  | ctccttcccc  | cacctcttgg  | gcgtctccca  | ccctgtggga  | acaccccact  | 103020 |
| tcccccttgt  | ccactcagcc  | tggctggggg  | cccagagtgg  | gagccggccc  | aggagcttcc  | 103080 |
| tgggaggctg  | ctgcgccttc  | ggaatgttta  | accccgact   | ccttttctcc  | aaaaatgcac  | 103140 |
| tggcctgggg  | ccctgtccaa  | gggtctcaga  | gtctttggag  | ggagtctctc  | cttcgcaagt  | 103200 |
| ggggagcaga  | tggctccttg  | ctccctggcc  | acaggcccca  | caaggcctcc  | agcatgagct  | 103260 |
| catgaggctg  | gaatgccact  | tgttttattg  | gggaaagggtc | tgcaccggga  | aaaaggccat  | 103320 |
| actcgaggtc  | cctgttcttc  | tgcagccctc  | gctatcttta  | ctcttgccct  | cctggtacct  | 103380 |
| tgcccccttga | tatatacccc  | tcatcttgaa  | atgtgagtgt  | ttcctgcctt  | ttggaggggga | 103440 |
| tacctagcct  | ctactctttc  | ttctgtacca  | tcttggcagg  | cttcctgggg  | gcaggggccc  | 103500 |
| accggtgggg  | gaagcagagc  | cccttggggg  | ctctctctct  | ggtcacagcc  | caggccagac  | 103560 |
| agacaggagg  | gcccagaggc  | agagtgacc   | cagtgtgtgt  | ccagccttcc  | cctcctgggg  | 103620 |
| atggggaggg  | caatctcaaa  | gctcaggcca  | gtgcgctgct  | tgaccagtgg  | aatgggggcc  | 103680 |
| ttatgggcct  | aggggatccc  | agtgagggcc  | ctgggttggg  | agetgctggg  | tctctggggg  | 103740 |
| cctctcagcc  | ttcatggcaa  | tgtcccttg   | ccttccctct  | tgtggattt   | ggacagtagg  | 103800 |
| gctgaaaatt  | ccaaacaaag  | agggtctctc  | aggaggggca  | ggggtgtagc  | caatggttta  | 103860 |
| aaatcgttca  | gaccttagtg  | ggtctcaggc  | tcccagccta  | aagagctgtg  | tgaccatgga  | 103920 |
| caatttcccc  | aagctctctg  | ggcttccgtt  | tgccccctct  | taaaatgagc  | atatcaaggc  | 103980 |
| tactgccttc  | ttagtttgca  | gcacagatat  | tatggcacia  | acagatgggg  | catggttatt  | 104040 |
| ctggaagcgt  | gtgaagagcg  | ggattgggaa  | gaggctgggg  | cagagcgtcc  | tgcagaagaa  | 104100 |
| gcacatgggg  | tggctcttaca | tctgggggac  | atcagagag   | tgaccactgc  | ccccccata   | 104160 |
| ccagaagtgg  | attccacagg  | agccagttag  | gctgaagggt  | caggccttcg  | tggcagggcc  | 104220 |
| ctgagaggga  | cagcagtgtg  | tccacagggt  | cacatgttct  | ggtcaacttt  | gcaaaagggt  | 104280 |
| ttcttttttg  | tgtttttttt  | tttttttttt  | tttttagagg  | ctcctgaaaa  | gcttcaggac  | 104340 |
| ccacaaactc  | tggaccatt   | tctgcctggg  | gggggtgggg  | gtggcccaga  | tcatccaggg  | 104400 |
| agggagggaa  | agagggaggt  | ggggtggaga  | aagctgaaat  | gacttccatg  | tgtgcgggct  | 104460 |
| cacgagatcc  | agatgtccaa  | accccagtg   | cttcttctgc  | ccacttgagg  | ggcaggggag  | 104520 |
| gcaggggcct  | ataggagtag  | tgacttgggtg | gttctgggga  | ccccagcaaa  | actagaagct  | 104580 |
| gtaatgtagg  | gagagacaaa  | agggctggga  | ggttcagggg  | ccctgtggag  | ggcggggaga  | 104640 |
| catggcactg  | accggctcct  | ccaggctgac  | ggtgcgccag  | ggttgtccat  | ccaggaccca  | 104700 |
| gtgcgggggtg | actggctgcc  | cagggatatg  | tcctggagta  | aagacagagc  | acagggtgag  | 104760 |

|             |             |             |            |             |             |        |
|-------------|-------------|-------------|------------|-------------|-------------|--------|
| ggggacctga  | ggaacacag   | ggcatgggac  | aaagcagagg | gaggggggta  | gaggacatcc  | 104820 |
| ccaggaggag  | actggaggcc  | ttttggggca  | gacttcacct | tcaacacgcg  | tgggctcagc  | 104880 |
| ctggagaagg  | agggacgccc  | gtgggcatcc  | ttggatctga | ggagctatca  | aggaggagga  | 104940 |
| aaagagaagg  | ctggaaagg   | acagctcagc  | tggggacacg | ggagtccctt  | gacctttgtc  | 105000 |
| ggggggcagg  | cttgggctcg  | gcgatcacaa  | ggaagaggcc | aaggccgcca  | gtgcagaggg  | 105060 |
| gaaggaaaga  | gcgcggcagc  | cttagggatt  | tttagatggg | cagcagatgc  | ctttagggtg  | 105120 |
| agagatgtac  | gaagagagga  | cacttgtgcc  | cccccatca  | tctgagaaaa  | acaacagcca  | 105180 |
| gatgttgcc   | tgcgaggtcc  | accttgccca  | gagctccctc | ggggactctg  | tcttggtggc  | 105240 |
| agggttttgg  | tacctggcc   | cagaaggccc  | ctctcatct  | cttcaagggg  | aggggacgct  | 105300 |
| tccggacgga  | gccttggtgc  | tccctggccg  | gggtgccta  | agggggctct  | aggaggaatc  | 105360 |
| ccagagccaa  | gcattactca  | gagggcgcc   | ggaatgttcc | cctggaatgc  | tcccagcccc  | 105420 |
| tccactggcc  | ccaaccactc  | tcacaggccc  | gccttgccag | agccaggccc  | caggcaccca  | 105480 |
| gagcctgcag  | cagccctcct  | tccccggta   | cccagtccca | gctcccagaa  | cagacagcct  | 105540 |
| ccccctcca   | cgcagccctg  | gcctcagtc   | tgctgggctg | atggctgcct  | gtggaagtga  | 105600 |
| ctcagctcct  | gctaggccac  | cccaactcct  | ttttctcct  | ccacctcttc  | tcccagacta  | 105660 |
| caaacatcaa  | agacccttcc  | tccaagaagc  | cctccttgat | tggatgagtg  | aattgccatc  | 105720 |
| aggcagatga  | gggcccagag  | gagctcgcca  | ccttggaag  | gaggctagag  | gggccagtgc  | 105780 |
| agggagggct  | ctgagtgat   | gtgggggag   | ggaaggagg  | gaggtctctc  | agcccagaga  | 105840 |
| gcacttaact  | gagagtagag  | aaccaagctt  | tgtgtcctc  | agccctctaa  | gggtttgggg  | 105900 |
| aagaggtagg  | gtggggcccg  | gcacagggtg  | gggtggggtg | cagtgtgggtg | tgtgggtgct  | 105960 |
| gtccacatgg  | ccttgccgcg  | acgtgctggc  | cacgggcacc | ctgaccccaa  | tgagggagag  | 106020 |
| aggggcagag  | ctggagctgg  | agctggagct  | ccggtgaccg | ggtgaatggg  | ggtggaaccc  | 106080 |
| gagggagcca  | ggctgggtatt | gggcacatag  | acgcccctct | cccaggggtc  | ccatcacctc  | 106140 |
| ccctgacccc  | aggatagggc  | tcagagggga  | gggagcagtg | gaccgcctgg  | ggccctcccc  | 106200 |
| tggggccaga  | acagaccagg  | ccctgtacc   | tgtttggtcc | ccacacagtg  | ctgtggaagc  | 106260 |
| caccgcccag  | tctgcatagc  | acagcccagc  | cccgcatgcc | ccctccctgg  | ttgccctccc  | 106320 |
| tgttcccggc  | caggcacttg  | ctgtgcagga  | ctggctaate | ctcccacccg  | cttgccagagg | 106380 |
| ttgttccagc  | cccatcttaa  | catctttgtt  | tggaggggtt | accccagga   | gacagctgca  | 106440 |
| gtctttccag  | agcactgcta  | aacagacacc  | ttctatctgg | agaggccctt  | ctctatctca  | 106500 |
| ccaaacaagg  | caacaatata  | aacaacatac  | acactgccct | gctgccctgg  | gaggaggggac | 106560 |
| gaggggtgag  | caggggtggag | gccacagcta  | gttctgcagc | ctgagagcaa  | agcaggggact | 106620 |
| ctgggggact  | cttgggcatg  | ggggcttcc   | agaggatgga | gccccgctga  | gtcctaagg   | 106680 |
| gtggaggagc  | aggagcgggt  | cacacgggtg  | cctgcggatg | gaagctggtt  | gtgagagcga  | 106740 |
| gaatccaggc  | agagggggct  | acggctatgg  | gctgggggct | gggggctggg  | ctgtccccga  | 106800 |
| gggggagggg  | gccatgccct  | ctgctttgcc  | agcggagtg  | cagccgggca  | gtgtggggcaa | 106860 |
| gtccggggccc | ggggccagcc  | caagcacact  | tgagcgtccc | tgggcaggtc  | ccacggagac  | 106920 |
| ccccccaaag  | agtccccacg  | ccctgaccta  | ctggccgtat | ggtgccgggg  | ccgtgagacc  | 106980 |
| ctccgcgcgc  | tgacccgagc  | tctgagcaga  | acccatcccc | gccaccacca  | cgcgcctag   | 107040 |
| cctgcccttc  | agggcgccac  | ccgcccgcgt  | cctcaccttg | aagcaccctg  | gcgcctggca  | 107100 |
| ctggccagag  | cagcagcagt  | agtagcagca  | gcagcaacgg | ggtccccoga  | gctctccggg  | 107160 |
| gcctccagcc  | catagctgtg  | agctcctcgg  | cctctaggca | gcggctcgca  | actccggctc  | 107220 |
| cgcaccaggct | ggattgcggc  | cgaaccgtgc  | ccggtgcagc | ctcaggccgc  | cgccttcgga  | 107280 |
| ccttcccgc   | cccacctccc  | accgcccgc   | ctcgtcccg  | cctccctcc   | ccgccaaacc  | 107340 |
| cgtcggagc   | ctggccagg   | gccccgacgg  | cgcgcgccat | gggggagccg  | ggtcgccact  | 107400 |
| ccgggaccgc  | cgcctctcga  | gggggtggag  | ctgggcggag | gaggggaatc  | gtgcggcccc  | 107460 |
| tccgatgacc  | ggcccgagcc  | gtccctcccc  | gtcggctca  | gagggcctct  | actcctgaga  | 107520 |
| ggaggagaga  | accgctggga  | aggttcttgg  | aggaccgcgg | cgtgggtggga | tgaggcggtg  | 107580 |
| ggcaaaggcc  | gcctctcgct  | gctgaagtgg  | gccccaggag | cgcgatcttc  | cgtgggtctcc | 107640 |
| tggggccgat  | ctctgtcccc  | tccctgtctac | cgtcctgcc  | ccgagggtgc  | cctggcgagg  | 107700 |
| gttgagtcgg  | gtcatccacc  | tgcactgggt  | gcccccaagg | ataggaaggt  | tcaggcaacc  | 107760 |
| ggctgccgct  | gtcttggggg  | cttcattgct  | gggcaaaggc | gatgcagcag  | acggagacaa  | 107820 |
| cctttcttcc  | ctggcggtgg  | ccagagggca  | gaattgcata | aaagctgcag  | actcccaggc  | 107880 |
| ctgggagacc  | ctttcggcct  | cagtaacatc  | tgtttcatgt | tttaaacttt  | tgttttccta  | 107940 |
| ctcggtgcaa  | atttggtatga | gatgttaact  | tttttttttt | tttttttttt  | gagatggagt  | 108000 |
| ctcctctgt   | cgcacggctg  | gagtgacgcg  | gcgcgatctt | ggctcactgc  | aacctccgac  | 108060 |
| tccctgggtc  | aagcgatttc  | cctgcctcag  | cctccagagt | agctgggact  | acaggcgcg   | 108120 |
| gctaccaccc  | ccagctaact  | tttgtatttt  | tagcagagac | gaggtttcac  | cattttggcc  | 108180 |
| aggatggtct  | caatctcctg  | atctcgtgat  | ccaccgcct  | cggcctctca  | aagcgctggg  | 108240 |

38/154

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| attacaggca  | tgagccaccg  | cgccccggccg | gagatgttaa  | cttttaagca  | aatctttttt  | 108300 |
| tttttttttt  | tttttgagac  | agagtttctc  | tcttgttacc  | cagactggag  | tgcaatggca  | 108360 |
| tgatctgggc  | tcactgcaac  | ctctgcctcc  | cagattcaag  | tgattcttct  | gcctcagcct  | 108420 |
| cccagtagc   | tggcattaca  | ggcattcgcc  | accacgcctg  | gctaattttg  | tatttttagt  | 108480 |
| agagatgggg  | tttctccatg  | ttggctcaggc | tggtctcgaa  | ctcccgacct  | caggtgatct  | 108540 |
| gcccgcctcg  | gcctcccaaa  | gcgctggaat  | tacaggcgtg  | agacaccgca  | cccagcctac  | 108600 |
| ttttaagtaa  | atctatttgt  | ttttgagaat  | ttggaatgta  | gtaatttggt  | tagtgaaagt  | 108660 |
| tcgagcagtg  | agagaaaacct | acattcacat  | atctcaaaaat | caaaaagtac  | agaaagcata  | 108720 |
| gggaaaagtc  | tccgtgctct  | tagccctcct  | caccaacagg  | aaaccaatat  | gattagtttc  | 108780 |
| tttcataggc  | tttttagatta | ttttttcaca  | ctcaagacaa  | tacagacata  | tttttttctc  | 108840 |
| ttattaacgt  | ttttctgcac  | tttgattttc  | tttttttttt  | tggtcgctta  | atacacctta  | 108900 |
| gatatacgtg  | cgttttagagg | gtccttggtg  | ttcttatgat  | tattatttag  | agacagggtc  | 108960 |
| tcactctgtc  | acccacgcta  | gaggacagtg  | gcctgatcat  | gcctcattgc  | agccttgaaa  | 109020 |
| tcctgggctc  | aaggatatcct | cccacctcag  | cctcctgagt  | agctggaact  | acaggcacac  | 109080 |
| ggcaccaggc  | ccagctaaaa  | tttttaattt  | ttctgtagac  | aggggggtctc | actttgtttc  | 109140 |
| ccaggctggg  | ctcaaaactcc | tggtccttggc | caggcgcagt  | gtctcatgcc  | tgtaatccca  | 109200 |
| gcactttggg  | aggccgaggc  | gggcagatca  | ctggagggtca | ggagttcaag  | accagtctgg  | 109260 |
| ccaacatggg  | gaaaccccat  | ctctactaaa  | aatacaaaaa  | ttagccgggc  | atgggtgggta | 109320 |
| gcgcctgtag  | ttccagctac  | ttgggaggct  | gaggcaggaa  | aatcgcttga  | actcagaagg  | 109380 |
| tgagggttgc  | agcgagccga  | gatcatgcc   | ttgcactcca  | gcctgggcaa  | caagagcgaa  | 109440 |
| actcgtctc   | aaaaaataaa  | aataaaaaata | aaaagaactc  | ctgatcttaa  | gtgatcctcc  | 109500 |
| tgccctagct  | tctcaaatcg  | ctggaattac  | aggagttagt  | caccacagct  | gtccagctac  | 109560 |
| gagattatta  | cttattatta  | ctactttgga  | ttttcaaact  | aacttcatta  | aggtataatt  | 109620 |
| tacacacaat  | aaaatgcact  | tattttaagt  | ggccagtaag  | atgagtttgc  | ataagtgtat  | 109680 |
| ataactacat  | aagcatcact  | ataatgcaga  | cacattccct  | cactcacaga  | aagagccctg  | 109740 |
| tgcccttcca  | gccaaacttc  | cccactccca  | accccagaca  | gccactgac   | tggtgttctc  | 109800 |
| tgtctataga  | taagttttgc  | ctgttctaga  | atttcatata  | aatggaatca  | tgccggcatgc | 109860 |
| actcttctgt  | gtctggcttc  | cttccctctt  | tccgatgttt  | ttgagattca  | tttacactat  | 109920 |
| tttgcatatc  | aatagtttgt  | tccttcgtat  | tgctgaatag  | tggtcgggtg  | tttgagggaa  | 109980 |
| ccacagtttc  | ctactccacc  | agtgcacat   | agggttatatt | tccagttagg  | ggctcttata  | 110040 |
| attggaacta  | tatttgcaca  | gagagagaga  | gagggaagaa  | gagggagaga  | gatattttatt | 110100 |
| atagcaattg  | gctcacgtga  | ttatggaggc  | caaaaagtct  | ccgaatctgc  | catctgcaag  | 110160 |
| ctggagaacg  | aggaaagcca  | gtggtgtgat  | tcagtttgag  | ttcaaaggcc  | tgagaaccag  | 110220 |
| gagcaccagt  | atggagggtg  | ctcgagctca  | gaacaagtgt  | gggacaggaa  | agcagagcag  | 110280 |
| caccccagag  | cagccctcca  | gcgacacctc  | ttcagtaaa   | caaggctgaa  | cacagagggg  | 110340 |
| ctggcttcag  | tgtggatgtc  | aggtacagaa  | ggcagctcga  | ggagctactc  | tgccgttctt  | 110400 |
| gcttactggt  | attcttacct  | cgaactggcc  | aactcctact  | taaactgcag  | gccatggctt  | 110460 |
| taatgtctgt  | tcattcagag  | gctgtccctt  | acccaaagcc  | agggttagcat | ccctgactgt  | 110520 |
| acactctctc  | ctgcaacacg  | tttcagaagg  | ccctgtagtc  | gtccacttcc  | ctgtctctct  | 110580 |
| ccccaaagctc | ctgagctcca  | tgtggtctgg  | gaatatgtgt  | gttgctcact  | tcctagcaca  | 110640 |
| gtcagtgtca  | ataactgact  | gtagagggga  | cacagtcgaa  | aagccacatg  | gggatcagag  | 110700 |
| tcattccttac | acagttgaca  | cctcccaaac  | ccagatgagc  | tgtgtccaag  | tgccaggtcag | 110760 |
| aggaattttc  | tgcggaagtc  | tctgagaaa   | ggtttattta  | cattttgagg  | ttgcagggga  | 110820 |
| ggagatgagg  | ccatcaaac   | aaagctgagg  | aagagggatc  | ctaggatgca  | ccgagcagct  | 110880 |
| ccggggggcgc | ctgacagcac  | ctgggaaaga  | tggcttctcc  | actggcttgt  | tgccgtcacc  | 110940 |
| ctccagaggg  | gcacagggaa  | atgtcctggg  | aaccaggcaa  | accagtgagc  | attaaccctt  | 111000 |
| agaagtgtct  | ggcatgggtg  | acaccacca   | tctgtaaa    | cgacttctcc  | caaggagtga  | 111060 |
| cgcagaacag  | gatgtctgag  | ggaggcactc  | cgactccagc  | cttcagagat  | cgccaggggtg | 111120 |
| gcacctgggtg | acgacaggct  | gatgcttggg  | tgccccagaa  | aaggctcatgt | gtgtgaatgg  | 111180 |
| gggccccaaa  | gccaacgctt  | catccctgac  | agcctgggtc  | atttagaggg  | gaactttttg  | 111240 |
| tccttgggca  | agggtgggtg  | aatttcaggt  | tcatagggca  | agggtatatt  | agctttaata  | 111300 |
| gatattgtca  | aacagttttc  | caaagtcatt  | gtacacactc  | tgtgattcta  | cttatgtaaa  | 111360 |
| gtttaaaaac  | aggcaaatca  | aatctatggt  | gttcgaagtc  | aagacagtag  | ttacccttgt  | 111420 |
| gggggctgca  | actgggtacag | agtgtaaagg  | gggactgtag  | gatgggtctat | ttcttgatct  | 111480 |
| gggtgtgttc  | gcttttggaa  | aagtccctga  | gttgcattta  | taatgtgtga  | acttttctgt  | 111540 |
| atgttacact  | ttaattgaat  | gtacaaaaag  | tctcaggagg  | cctcagacca  | ctggaagcgg  | 111600 |
| acacaactaa  | cccctctgag  | agcctccaat  | ccaagatgga  | catatgtccc  | cttggaagta  | 111660 |
| tgccagaagca | ggtgaagact  | cctaagccgg  | atattcccaa  | atccccccag  | tagccgcagc  | 111720 |



|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ttcagcagct  | gcttatggtc  | ctccctacac  | cctctcttcc  | ccagacagcc  | cccaaacatc  | 111780 |
| tggttgcat   | tgacttgctc  | tctccctgtc  | ccacctctgg  | atttagtcca  | tggttctccac | 111840 |
| cctccccact  | gtcagcaatg  | tagacaagac  | aaacgcttag  | ttcacgtgcc  | cacctactgc  | 111900 |
| gtgccatgca  | cggggctggt  | cattgtgggt  | ggcaaagtgt  | agcaacacac  | gaagcctcaa  | 111960 |
| ggagcagaaa  | gggacacaaa  | tcacttcagc  | gtaaggtaat  | ttgtgataaa  | tgtcatgtaa  | 112020 |
| cttgccagccc | ctggcccccct | cctacagatg  | gtgtctaaga  | ataaacccca  | ctaactgtgt  | 112080 |
| actcctctgt  | tctagcccag  | ctgtttgggt  | tgcaagaaag  | agactcactc  | cagttgctgc  | 112140 |
| ataggatgga  | gtttttattg  | gaggacattc  | tggaagggtc  | cccagcaaga  | gtctggcaat  | 112200 |
| ggagcatgaa  | aatgaatgag  | cctggaatga  | gggagggtgc  | ggcctcgggt  | acacaaagt   | 112260 |
| cacccggccc  | ccaactgcct  | cccagcgtgt  | tagctcctgt  | ggcaactccc  | cacttctctc  | 112320 |
| tctgtttttc  | aacccaaatc  | ctagagcaga  | gggtctctcg  | ctcctcacac  | catctccacc  | 112380 |
| aggctgcagg  | gggagtcact  | agcccactca  | agagctctct  | cctgttgggt  | cctgattcgc  | 112440 |
| aggggcacgt  | cattccttct  | tggttaactca | ttctctttta  | cagcccgaca  | caggtctccc  | 112500 |
| agggaaagga  | caggagctca  | cacagtcatg  | aacagagatg  | catctggagc  | aggaataagg  | 112560 |
| atgctgacaa  | catctgtgtc  | tgccctccac  | tcttgtaagt  | gccagtaagg  | tagtaccag   | 112620 |
| tccctggggg  | taggggggggt | ggtgggtggga | atgagggcat  | ctcttctttc  | agggggccaaa | 112680 |
| tctggcacia  | ggatgtctgc  | ttaggcaact  | ccactgcctg  | gcatgctctc  | tccctaggaa  | 112740 |
| aacaccaaac  | cttttttatt  | tctcagctc   | tagtgtgaga  | gtgatcacct  | cttccaggaa  | 112800 |
| gccacaaccc  | agaatggcca  | ggggcacacc  | ctggccctca  | gtagatgggc  | actgagacaa  | 112860 |
| aacatggcct  | gctggtggca  | ttcccaacaa  | tgtcaaaagt  | ctcagggatg  | agctcacagt  | 112920 |
| tgggagtctc  | ttggaagtcc  | aattccaaat  | gtcctctgga  | ctcaaaataa  | aaactacatt  | 112980 |
| tcccagtctc  | ccttgtagct  | tgtagtagcc  | atgtgacaat  | actccagcca  | atgggaagtg  | 113040 |
| agtggaaaag  | tctgtgtcag  | cttctgagtt  | gcgtccacca  | tgggcatggg  | tgtgcttcta  | 113100 |
| cctccactct  | tcttcttctc  | ggctggaggg  | catgtggaca  | aagcggggcc  | atggtgagcc  | 113160 |
| tcacaatcaa  | aggcatcatt  | ttagggatag  | agaaaaggaa  | gatagaagga  | tcctgagccc  | 113220 |
| caacatcata  | aaccatcgta  | acagccagac  | tagcgtggga  | gagaaaaata  | aaattctatc  | 113280 |
| atgttgaagt  | cactgtatct  | ggtctcttgt  | tagagcaggc  | tgactgatac  | tctgttgaat  | 113340 |
| acgaatgcc   | ttggtgagcc  | tctgaggatg  | caggatgctg  | caatccaacc  | aaagtccagg  | 113400 |
| gtgactttcg  | aggagaagga  | tgtgaaaggg  | cagggcttcc  | tttgaggggag | aaatggagag  | 113460 |
| aaggaggagg  | atccagagaa  | gagagagcga  | gatcagctca  | gtgctcttca  | gtgtcctcca  | 113520 |
| gggtcaacac  | cctcctgggc  | agacctctca  | ccaccatttt  | ttgggggggc  | aagaggcttg  | 113580 |
| gggcccaggc  | ataaaaagtc  | agatgtcaca  | gagctgtttt  | cgtagaggcg  | ggcaggactc  | 113640 |
| aacactgcct  | cattcacatt  | cataggctgg  | agtcatcaca  | gagtctggcc  | actttctcct  | 113700 |
| ccggagggca  | ggcaaacacca | ctggtcagcc  | caggggtggg  | gcacaggttg  | aggtctcaca  | 113760 |
| tgtggctgca  | tcaggatcaa  | tgagcttccc  | acagagaaaa  | ccctggtcac  | aggaggcctc  | 113820 |
| tgggagccag  | ggcaggcccc  | caggccccc   | gcctcagggt  | gggtcccaag  | aggattgggt  | 113880 |
| ggagtgtgca  | tggctggaag  | tgtgacttgg  | agcctggcc   | taccaggtga  | atcagctgca  | 113940 |
| gaagggacac  | accacagggg  | aggagaaact  | gactttcact  | tctgcccagg  | tgacgggcag  | 114000 |
| gcttgggagt  | ggggaagagg  | cctcagaaca  | aaggtggggg  | agtcagaaca  | tacgtgtctt  | 114060 |
| ggaaatggag  | gggatgctgc  | caagtccaga  | aagtctcccc  | tagtctctca  | agccagccgc  | 114120 |
| cagctgacac  | agggattccc  | tctgccccga  | gcagaacagg  | gctctcctat  | ctcctgaggg  | 114180 |
| gctcctgggt  | ccctgcacac  | tctccccaac  | ctccccttgg  | tcccaggcac  | catcctctaa  | 114240 |
| atgacagcag  | tctccagatg  | ccagttggca  | ctgtgggtta  | gctgggtatt  | atatctgccc  | 114300 |
| gtgtgtgttg  | aatggatacc  | tgcgtggtct  | ctttctgaaa  | agccatctcc  | accgaattct  | 114360 |
| cgcccatgct  | ctgcttacaa  | acacgcctcc  | ttctgcaagg  | cccggagtgg  | actcagagag  | 114420 |
| cctctccccc  | ccccacctcc  | tgcccccat   | cctctcctgc  | tccccacagg  | ctggctccag  | 114480 |
| accacaagag  | cgtgggaacc  | tcatggagag  | tgggtgcttc  | ctctgtcttc  | ctccctagcc  | 114540 |
| cttgcttttag | cagagccagg  | gcagagagga  | ggaaagaagg  | aaactacagc  | aaggaggatt  | 114600 |
| tagggacaat  | tctagaaggg  | atttccaatt  | tggaggggca  | gtctgggaag  | taggcagcct  | 114660 |
| gtgaacttaa  | taggaggtcc  | taaaggacct  | gggggaattt  | ggggagaaga  | cgaggtgggc  | 114720 |
| tgccttaccc  | caatcttctg  | gcccactccc  | agctccagac  | ccacctccag  | gtgtagcagg  | 114780 |
| ccccagggcc  | caacagcagg  | agcaggaggc  | ccaccagcag  | tcccaggacc  | cagagcagct  | 114840 |
| gctggaactg  | atgcaggcgg  | tgcagggtct  | gaacacagag  | cgggactcag  | agcagccaca  | 114900 |
| gctgcaggcc  | cccagtgtct  | ccttccaagg  | ggaccacccc  | aagatgacac  | cttctaactt  | 114960 |
| ttgcttttat  | gctgaagctc  | tccacacaag  | ggtgccccga  | gcaatcagtt  | tagagtgcag  | 115020 |
| aggcaaatca  | tgtctgcagca | ggagggatag  | agggaaaggaa | gcctttgggc  | ctaagagctg  | 115080 |
| ggcttttgtg  | cagatgggcc  | tgggagatgg  | agccccctcc  | cctcacctct  | gaccccaag   | 115140 |
| taggtggagg  | tagaggcaga  | gcccaggaca  | tttgaggcag  | aacagatgta  | ttcccccttg  | 115200 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| tcgctggggc  | tcagcgccctc | gatgtccacc  | ctcagggcat  | tgggggaagc  | caggacatgg  | 115260 |
| atgtgtggct  | ctgcaggagc  | accctggggc  | tgactggagg  | ccaccagtcg  | actgccaaag  | 115320 |
| tggagagtca  | ggctggcgag  | cggctcgctg  | tccactcggc  | aatccaggat  | gccccggagg  | 115380 |
| ccaccctcag  | gctccacgaa  | gacctcatg   | gtgggcgtct  | tgggagggtc  | tgtggggagg  | 115440 |
| aagggagtgt  | ggtgattgca  | gaaaatgacc  | acaaatttct  | ccctccctgt  | acccatgcac  | 115500 |
| agtgactttg  | ctgctcctcc  | cattaagagg  | cagaacctat  | ttcctcactc  | cttgaatctg  | 115560 |
| tgactcagtt  | tgaccaattg  | aagaaagaag  | tgatgtttgt  | caacttcaga  | gccagacctc  | 115620 |
| aagaggcctt  | gtagcttctg  | ctctacttgt  | tggaaacaaa  | caatgtggga  | agcccaggcc  | 115680 |
| agcctgctgg  | gcacatgtgg  | cccaactggc  | agccaggagc  | gtgcagtcac  | ctgagaccac  | 115740 |
| tgggcaactg  | cagccacatg  | aggtgtgaac  | aacagaagaa  | ctccccagct  | gagcccagcc  | 115800 |
| cacagaattg  | agcagataac  | atgacttttg  | tttgaagcag  | taagttttgt  | tgtggtcaat  | 115860 |
| agcaatagct  | gactgataca  | gtgtaggcct  | tgggtgaaggc | tggagtggga  | gaccaagact  | 115920 |
| gatgagggct  | tatcgtgtgc  | cagacacact  | cagcacgctt  | tattttctttt | ggtttttttct | 115980 |
| cttcttttct  | tttctttttt  | tttttttttt  | ttttgagaca  | gagtctcagt  | ctgtcaccca  | 116040 |
| ggctggagtg  | cgttgtgggtg | atcttggtct  | actgcaacct  | ccacctccca  | ggctcaagtg  | 116100 |
| agtctcctgc  | ctcagcctcc  | tgaatagctg  | ggattacagg  | tgtgtgcctt  | catgccgtgc  | 116160 |
| taattcttgt  | attttttagta | gagacggagt  | ttcgccatat  | tgaccaggct  | agtcttgaa   | 116220 |
| tcttgacctc  | aagtgatccg  | cccacctcag  | cctctcaaag  | tgtggggatt  | acaggcataa  | 116280 |
| gccccctgcac | cagccattta  | tttcttatac  | aacttgacag  | atgaggcgaa  | tgatgtatct  | 116340 |
| ccatttttac  | aaggtcaccc  | aaccagacag  | cgggagagcc  | aggattccaa  | cccaggactc  | 116400 |
| aatatccgtg  | aagccttcac  | tcttggcaga  | cagaagagga  | ggaagaactg  | aggtgggctg  | 116460 |
| gtgagatgtg  | gggagactaa  | ggcggaggag  | ggggttcaca  | ctcacagagc  | acacggagca  | 116520 |
| tgactggagc  | agaggcagca  | gccccagtgg  | ggagctcagc  | caggcagtg   | tacatcccag  | 116580 |
| cttgagcacg  | agccacgtgg  | gtgaaggcga  | gagtgggcac  | aggctccgcg  | tgcagccgcc  | 116640 |
| ggtcattcca  | gaacctatga  | aaggtggagt  | tgcccacagg  | cccaggggcca | cccaggaggc  | 116700 |
| ggcagctgag  | gttcaggggca | gcgcctcag   | gcacgtccag  | gccaggctct  | gccaccacgc  | 116760 |
| gtgcacctgc  | gggcccaggga | tagagagatg  | attggggatc  | tgtaggcctt  | gggggctgga  | 116820 |
| gcctctgggt  | gggacctctg  | aggggcctct  | gcacattgtg  | ggaaggctct  | agtctgactt  | 116880 |
| ggtatagggc  | tttccaaggt  | ggaacctatg  | ccctccagct  | gggagctgtg  | ccccctccac  | 116940 |
| cagattaggc  | ttctcaaggc  | agagctttct  | cagaccagac  | agcccgtctc  | agtgggagct  | 117000 |
| ccagccctct  | tctgtggactt | ggacctgaga  | gcacagagc   | ccctcctcct  | ccctctctct  | 117060 |
| ccactcacct  | tctacctgca  | accgcccgat  | ggtgctgatt  | gagcccagca  | agttttgggc  | 117120 |
| tgtgcaaaca  | taggtgtcat  | cacctgcagg  | cacatcttgc  | acctgcagcc  | gtagggcgtt  | 117180 |
| tccggccacc  | tggacatggc  | ctgtccctga  | tgccaagctg  | tggaccccgc  | tgtcgtggc   | 117240 |
| cagcaccttg  | ccatcatgag  | ataggggccag | ctcagcaggt  | ggctcactgt  | ccacagtga   | 117300 |
| ctgtatcaca  | gccatggatc  | tggccctgga  | gtcccggaa   | gaggacagga  | cagcgtcctg  | 117360 |
| aggggcatct  | gtgggcaggc  | agggcacaga  | tggggacctt  | gcttaggcac  | cctgtactgc  | 117420 |
| tcattgccta  | gctgcctggg  | gttggtcagg  | ctgaggcctc  | tggaccaatg  | gccacttctc  | 117480 |
| agaagtgcaca | ccttcccagg  | agtcctgtgg  | gcagagtgtc  | ccaggcaggg  | cttataggac  | 117540 |
| tgtctctttc  | tctagcacct  | atgtcctctt  | tccccaactg  | cccatctcgt  | ggcccactgc  | 117600 |
| agtccttttg  | ccactcccac  | cagagcccag  | aagcacctct  | caccaccctc  | ctcctgggca  | 117660 |
| gccttgggca  | aggaccctct  | gctcacagag  | gacttgacag  | gcagcaggac  | gggagctgcg  | 117720 |
| ggtgccttg   | gcacctctgg  | cctggcaaga  | gtaggcgcct  | gcacagagcc  | gcgtggccac  | 117780 |
| caggaatgag  | agtgaggcag  | ctggaccctc  | ctgcagccaa  | cgaccgttgt  | ggtaccaagt  | 117840 |
| atagagtgtg  | ggtgcgtggg  | cagcagggtc  | cgcacaggct  | actgtgatgg  | gggcaccttc  | 117900 |
| aggcacggca  | gcctccggag  | ccaggatcac  | ccgcacacct  | gtgccaagga  | ggccaggatc  | 117960 |
| agccggggcc  | agccggacac  | cacagtgggc  | ttccatccca  | gtgggcttgg  | cctaggccct  | 118020 |
| gccttaccct  | ccagccgcag  | ctccagggac  | gtgttgccct  | ggcccagagg  | gctgcgggca  | 118080 |
| gagcagctgt  | agaaaccctc  | atccctgggc  | tgtggccctc  | gcagctccag  | gcgcagggtg  | 118140 |
| ttggggacag  | aggctgctgt  | cgaggaggcc  | aagaggcgac  | cggcgtggct  | gagggccagc  | 118200 |
| tgggcccggc  | ggcggtgtgc  | cacagtgcac  | agtaccaggg  | ccagctgccc  | gccatggctc  | 118260 |
| tccaggagg   | aggtcaggcg  | caggttgccg  | ggcgctctg   | cagggcatga  | gaggtttaca  | 118320 |
| cggagtca    | ggcggcagcc  | tgcaggaggc  | cagtttgacg  | gtggcattca  | aactcaggag  | 118380 |
| ggccctggct  | ccccaccctc  | atggctccca  | accaccagct  | ctgaggcact  | gctcctctgc  | 118440 |
| ccagacagga  | ctcaccccag  | cgccccctac  | tcttaatgct  | ccttgcccag  | tgtccacta   | 118500 |
| gctactgggt  | accgagttcc  | ccactggaca  | cctgtcgggt  | tccggccatg  | ccgtgatccc  | 118560 |
| tgtgggcttc  | tgtcattcct  | gtgagtcaca  | ccctgcaccc  | agccagactc  | acagaggacg  | 118620 |
| tccaagggtga | taggtctgga  | gaggcggggt  | gcccagaccag | gggggcccac  | accgcagcgg  | 118680 |

|            |             |             |             |             |             |        |
|------------|-------------|-------------|-------------|-------------|-------------|--------|
| taggaggtgg | catccctgac  | tgtgacgttg  | ggcaggggga  | tggagtgggc  | atccaggcgc  | 118740 |
| tgctgcccac | cctggtacca  | tgtgtaggtg  | agctggggccg | ggtgagtggg  | ccacacaagg  | 118800 |
| caggtcaggt | tcaccagctg  | ccctcccg    | acggtagccc  | cgggccacac  | ctgcacattc  | 118860 |
| acagctgggg | agagggaggg  | cacaggacac  | cagttagggg  | cttgaggctg  | tgtacagcag  | 118920 |
| gccacctgtc | tccatgtggg  | tgagctactc  | ctactgctgg  | ggagcccctt  | ggcctttggg  | 118980 |
| agccttgggt | ggcccaacta  | taaatgtggc  | ttagaactca  | agcttgggaa  | tagcccaactg | 119040 |
| gctcctgagt | ggtcccaagt  | agagttccac  | agagtcctgg  | aaagccctga  | gcccctccct  | 119100 |
| gaccctggct | gggggcaaat  | ggggagacca  | cagacctccc  | ctcccctgct  | gcactttctc  | 119160 |
| ctgaaattcc | ctgtttagtt  | ttatccactg  | ggcatacttg  | taaaatttgg  | tgtttataaa  | 119220 |
| aggttctcct | aagaaactca  | agtctggagt  | ccaccatcgt  | taggaagggg  | gggaatgggg  | 119280 |
| actattcccc | tgccccaggg  | cttctagaac  | cctgtcccac  | ctcctctcct  | cttttaaagg  | 119340 |
| acacacacac | acacacacac  | acacacacac  | acacacacac  | acactgacct  | tgagcgctga  | 119400 |
| agtcagctga | ggccgaggcc  | tggcccaggg  | tgttgagggc  | ctcacagata  | tagacaccct  | 119460 |
| catcctccag | catagccccg  | tggatctcca  | gacgcagtgt  | gttggggggc  | acagccacat  | 119520 |
| gcagcctggg | agagctgcct  | tcgggtcccc  | ccacaccttg  | taggggtggag | gccacaaggc  | 119580 |
| gatccccgtg | gagcagccgc  | agctggggccg | gagggtcact  | gtccacacgg  | cacaggagga  | 119640 |
| ggcccagtcg | tccaggacct  | gtgtccatca  | gggtagttag  | tgtgacgtgg  | cgtggggcat  | 119700 |
| ctacacaggg | tggggagtg   | tcaggaccca  | gccagggggg  | tccctcatcc  | agggtcttgt  | 119760 |
| catgtgacac | agcccaggac  | tgggggactg  | cattccttcc  | cccacacctg  | gatgggacaa  | 119820 |
| aagtccttac | aggacaagt   | gaggtgtag   | ggtgcagcta  | ggctcgtggg  | ggctgagcct  | 119880 |
| ggggcctggg | cttgccaatg  | ataggcccc   | ccttgtgtca  | aagttatggc  | tgcaaaaggc  | 119940 |
| agcgtggccg | aggctgactc  | ctggaggggc  | tggccatccc  | gataccaacg  | atatgaggtc  | 120000 |
| ccctctggga | ctcctgtggg  | tacctggcag  | ctcaggacca  | cagcctggcc  | ctcttgagac  | 120060 |
| tcaggtgatg | gtgacacctg  | gacccaggct  | cctgcagggg  | aaaaccaaga  | gcaggtgagg  | 120120 |
| gctctccacc | acacctctca  | cagtctggga  | ccatgtgcgt  | gtccacctag  | gactacacaa  | 120180 |
| cccaacccca | tgtgctggag  | ccgcagcccc  | ttccagacca  | ccgccagggc  | ccactcagcc  | 120240 |
| ctatgccttg | gcccctcctg  | ctcccactc   | tcctgcacag  | gctccatcct  | ggcatttacc  | 120300 |
| tcccaccacc | tgggtgggatg | gttttactgt  | gtcagtcata  | tctcgcttac  | caaagcgtaa  | 120360 |
| gccttatggg | cactggactt  | gtgtgacctg  | taaccccagt  | tccaggcact  | gtgcctgcac  | 120420 |
| gtgttagctc | ctgacaaatg  | ttggttcaac  | ccataaagta  | ctgaaaaggg  | agggatctag  | 120480 |
| atcatcccag | ggccattgtc  | ctaactagga  | tctggactca  | gcatgggtga  | ccgagggcct  | 120540 |
| ggaagggaa  | gttagcagct  | atgttatctt  | gagcgtcttc  | gccaggccag  | ccctgcaccc  | 120600 |
| agactccagg | ccacaagagc  | acaggacctg  | gggaccaggc  | aaagggcagc  | tctcagtggt  | 120660 |
| cccctgggtg | tgaagcagga  | gaaccccagg  | ttgcagggag  | tgagatggag  | ggacagctgg  | 120720 |
| aatgctaagc | aagaacacag  | accatgcctg  | agaccacagc  | tagcaggggtc | acgcaactgg  | 120780 |
| ggaaaagctc | ctaaaactca  | aaccttcaat  | ttcttcctgc  | gaaataggta  | tgctaagtaa  | 120840 |
| gaaaagatac | acagaaccaa  | ggctcaaagt  | aaatgtccca  | aaattgggtg  | tgtcggtcac  | 120900 |
| cgtggtggaa | gaaacactca  | tgttcctttg  | accttcgcac  | tatgggcatt  | atcttttcctt | 120960 |
| tttcctgggt | tatttggtta  | gggagaattt  | ttttttttta  | agacagagtc  | tcactctgtc  | 121020 |
| atccaggcta | gagtgaaagta | caattttggc  | tcactgcaac  | ctccatctcc  | caggttcaaa  | 121080 |
| cgattcgctt | gccttagcct  | cccagtaaac  | tgggattaca  | aatgcccacc  | accacaccta  | 121140 |
| gctaattttt | gtattttttag | tagagtgttg  | gccaggctgg  | tcttgaactc  | ctgacctcaa  | 121200 |
| gtgatctgcc | cgctcagcc   | tcccacagtg  | ctgggattac  | aggcatgagc  | catggtgccc  | 121260 |
| ggcctgaggg | aggattttta  | tagagactat  | atataatatac | acacacacac  | acacacacac  | 121320 |
| acacacacac | acacacacac  | acgtatgtgt  | atataaataa  | agaatatgaa  | tatatattgc  | 121380 |
| ataagaatac | acataatgta  | tatttgtatg  | tgtatatact  | catacataaa  | catatatgca  | 121440 |
| taatataaat | atacatatgt  | atgagaatgt  | ttactatgta  | acaaaattat  | atgatcgcaa  | 121500 |
| aacaagttat | ttcagaaagt  | gctcccacca  | agcaagccac  | tttgtgaatt  | gccattgcct  | 121560 |
| cctagtggct | gtcgcagtta  | ttacagggtga | aaatatctag  | ctggaggcaa  | agggaggcct  | 121620 |
| tctgctggct | ggctggaaac  | tgcctttacc  | aggactgaaa  | aagacatatt  | tctgcacaaa  | 121680 |
| attcacagca | cccagcccaa  | agtctgacca  | gagtagtccc  | cagggcatca  | gagtctacca  | 121740 |
| tatcttcaga | agactgacac  | tcacctcgga  | cctggaagaa  | gagtgaggtg  | ttggatgatc  | 121800 |
| caagaacatt | tgtggcctca  | cagcggtagc  | tgccagagtc  | cccaaggccc  | agttctcgga  | 121860 |
| cctctaactt | cagggagttg  | gcctcagcct  | tagcctggaa  | ccgacctagg  | gatgggacct  | 121920 |
| ggggacccag | gctgggtggc  | aggaggtgct  | ccccatggaa  | caaggccagc  | aaggccaggg  | 121980 |
| ggcggctgtc | cacagtgcag  | atgaacagag  | ccatgtggcc  | ctggcccatg  | tctaggaggg  | 122040 |
| ctgacagctt | tggacgggtc  | gggggatctg  | caggaacaga  | gggagctgag  | gccacccagc  | 122100 |
| ctagctcacc | acattacaac  | tgccacacta  | tgtcccccc   | ctcctgccac  | acacacagcc  | 122160 |

|             |              |             |            |             |             |        |
|-------------|--------------|-------------|------------|-------------|-------------|--------|
| taagccatgc  | tcttttccctcc | ccaaaccccca | gcccgccccc | tgctttctccc | cagaccacccc | 122220 |
| ctccaccctc  | caagtcccca   | ggctgcccac  | gccagtccac | cacagaggac  | actcaggagc  | 122280 |
| acgggagtg   | agagctgggc   | accagcctca  | gtcaggatgc | ggcaggcgta  | aagggcagca  | 122340 |
| tcagttctgg  | ccacgggcag   | cagtgtcacg  | gtctccagg  | gacctggggc  | ccacagcacc  | 122400 |
| ccatttcgga  | accaggagaa   | gttagcagg   | ctgccagcag | cttcccggct  | cacgttgcaa  | 122460 |
| gtcaagttgg  | cttctgtgcc   | ctcctgaagt  | gtgtgtgatg | gtgcaatggc  | caggacagtg  | 122520 |
| gctggagagc  | aggcgggcaca  | gcttactgac  | cacccccggc | ccccaggagc  | caggggtcca  | 122580 |
| gccgcccagc  | ctggaaggag   | caggaaggag  | gccccctggg | gtgcccacag  | ggttggggag  | 122640 |
| aaaagcaagc  | tagctcacag   | ggcagcctag  | gagaagtggg | ccctggggac  | tgtggggggc  | 122700 |
| ctgggcagag  | agggttgccag  | tacagggaag  | gaggacaagg | cagcccaact  | gtcccagctg  | 122760 |
| gggagtcctt  | cctcagagac   | cagctgtgtt  | gcctatcccc | ggcctgaaca  | gagatcatgg  | 122820 |
| gaaggagcag  | ctccgctctt   | cctgaatgtg  | gaggagggca | gcgaaggggc  | cccactgctg  | 122880 |
| cacagagtgg  | ttttgcctga   | ttagatcctc  | ctcggagcag | aacagtccag  | agctccagcc  | 122940 |
| cctgccccca  | ggccacccat   | tccctgcctg  | actcaccctg | gccattgaag  | gtggctgagg  | 123000 |
| tggaggcggt  | gcccagggca   | ttgctggcct  | cacagaggta | caagccctcc  | tcttccagca  | 123060 |
| aagggttgtg  | aatctccaca   | cgcagcaagt  | tgggggcttt | ggtgaccttc  | atgcgtgggg  | 123120 |
| aacagccccc  | acaggtgctg   | cagccacccc  | ctgatggcag | ggaagtggcc  | acaacacggt  | 123180 |
| ccttgtggag  | cagctgcagc   | ctggcggggg  | ggtcgctgtc | cacacggcac  | aaaaggaggc  | 123240 |
| ctcccgctcc  | agccccggcc   | ccagcggcat  | caaggtccag | cctgggtggg  | aatgttgggt  | 123300 |
| gtcagggggg  | gtctgcaggg   | aggaagaaca  | tgggcactca | tcccacggat  | gctccagggc  | 123360 |
| cccacaagcc  | tggctaggct   | cccagaatac  | actggacaaa | ggcagatccc  | gaaggctgcc  | 123420 |
| ccaagcagct  | gtgcagactg   | tgcactgcac  | aagggcagcc | agaccagggt  | gggtggaact  | 123480 |
| caagctaggg  | tcatgtctcc   | caagctaagc  | tgtgccctga | tgcctactga  | gtcttcccag  | 123540 |
| aagaaggagc  | accatcttct   | agctcagaca  | aaggtgctgt | atgggctggc  | tgccaccttg  | 123600 |
| cataagtccc  | atctgctgct   | gagcagggtg  | cccacctagc | aatccctggg  | ggtcatgctg  | 123660 |
| tctgtcccca  | ccctgcaccc   | aatgccttac  | cacggagcct | cccctcaaga  | tctgccctgc  | 123720 |
| agcttatctg  | tgatctccaa   | gccagctcct  | gtgatactag | cgaatgcata  | gaatggcctc  | 123780 |
| ttcctgaagg  | cttctggaga   | cagttggcca  | aaagagagca | atcagccaag  | gccccactgt  | 123840 |
| gatctcactc  | agacaccaga   | aaactttgca  | gaggcagctg | ccctctgaca  | gttccgggac  | 123900 |
| gtggctccgg  | cacacatgtg   | tcaggggaag  | cctctggaat | caggcccttg  | gtcttacacg  | 123960 |
| gagcctcccc  | gcttggcgga   | tgcagagggg  | ctgtctgtcc | ctgtgctgag  | gggttttgcc  | 124020 |
| ttgggtgtcta | tgggagccca   | gcacagcctg  | tggggcatgt | gcacaaatag  | aggagtgtgg  | 124080 |
| tgtcccagg   | ctcaccatg    | aggacaaagc  | atggatgtgg | tgcccatgct  | gagcttgagg  | 124140 |
| ctgggacaga  | ggagcccaca   | gggctggcat  | tggagggtaa | agacatgggg  | aggtagagaa  | 124200 |
| ggccccagcc  | tcatttcccta  | aacttcagcc  | ccacgtggaa | agaccactg   | tggccaggtc  | 124260 |
| tgtgaccag   | ccctcccgat   | atagaccttg  | gggcagggca | ggacagggct  | ggggggccta  | 124320 |
| gaggaggtgg  | ggtggctggg   | gggtcccggc  | aggaggctgc | aacagggtgg  | ggctgctcac  | 124380 |
| agagcacagt  | gagaacagct   | ggcgaagagg  | ggccactggc | actgtggccg  | tcccggggcc  | 124440 |
| ggcagtggtg  | tgagccggcg   | tcagtgtctg  | aggccgcggg | gagcaggagg  | ctgtgtccgg  | 124500 |
| gacctcgtg   | aagcagggct   | ccattcaggt  | accaggagaa | gcgggcatca  | ggtgtggggc  | 124560 |
| ttaggcgct   | tctgcagctc   | agtgtcactg  | cctgtccttc | caccacctcg  | gctgccgggc  | 124620 |
| tgatgaggag  | acgggcggct   | gcggggagag  | gaagaggctg | ggaagggtcc  | ctcctctcaa  | 124680 |
| ccccacatgc  | tgccttatat   | agaaagcctt  | ccagggttct | cctttcccca  | cactactgta  | 124740 |
| ggtagctctg  | ggctttgtgg   | tgcatacagga | gggtttgccc | tttaattgcca | aatcagatct  | 124800 |
| attggtagta  | gatggctgca   | gcatggttga  | agattcagag | ccagagccca  | ggcttatgtc  | 124860 |
| caacacctgg  | cttggctggg   | catggtagct  | catgcctgta | atcttagcac  | tttgggagac  | 124920 |
| tgaggcagaa  | ggatcgtttg   | aggccaggag  | ttccagacca | gcctaggcaa  | catagtgaga  | 124980 |
| ctccatctca  | aacaattttt   | ttttttttga  | gactgagtct | cactctgtct  | cccaggctgg  | 125040 |
| agtgcagtgg  | tgtgatcctg   | gctcactgca  | acctctgcct | cccaggttca  | agcgattctc  | 125100 |
| ctgcctcagc  | ctcccaagta   | gctgggatta  | caggtgtgcc | accacaccca  | gctaatgtta  | 125160 |
| tacatgtagt  | agagataggg   | tttcaccatg  | ttagccaggc | agatctcgaa  | ctcccgacct  | 125220 |
| ctgggtgatcc | acccgcctca   | gcctcccaaa  | gtgctgggat | tacagggtgtg | agccactgtg  | 125280 |
| gccagctttt  | tttttttttt   | tttttttttt  | ttttgagacg | gagtcttgct  | ccgtcagcca  | 125340 |
| ggctggagtg  | cagtggcgca   | atctgagctc  | gctacaacct | ctgcctccca  | ggttcaagca  | 125400 |
| attatcctgc  | ctcagcctcc   | ctagtagctg  | ggaccacagg | tgtgcgccac  | cacacccggc  | 125460 |
| taatttttgt  | atcttttagt   | gagatgggg   | ttcaccacgt | tggccaggct  | gatctcaagc  | 125520 |
| tcctgacctc  | aggtgatctg   | cctgcctcgg  | cctcccaag  | tgctggaatt  | acaggcatga  | 125580 |
| gccaccatgc  | ctggccacaa   | tttttttttt  | ttaattagct | gggtgtgggtg | acatggggccg | 125640 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| tagtctcagc  | tattttgggag | gctgagatgg  | gaggatggct  | tgagcccagg  | agttttgaggc | 125700 |
| tgacgtgagc  | catgaacata  | ccattgcact  | ccggcctggg  | caacagagca  | acaccctatc  | 125760 |
| tcaaaaaaca  | aaaagaaaaa  | cctggcttga  | tcaattagct  | accatgccct  | caggaggagg  | 125820 |
| gaaggacagt  | gcacataccg  | aaagtggaa   | gaccgtactt  | ttcttttttc  | tttttctttt  | 125880 |
| tttttttttt  | ttttgagaca  | gagtctcact  | cttgtcacc   | aggctggagt  | gcagtggcgc  | 125940 |
| tatcttggcc  | cactacaact  | tccacctcct  | gggttcaagc  | gattctcctg  | ccttagcctc  | 126000 |
| ccaggtagct  | gggactacag  | gaactcacca  | ccatgccag   | ctaatttttg  | tatttttagt  | 126060 |
| agagatggga  | tttcaccatg  | ttggccagga  | tggtctcgat  | ctcttgacct  | cgtgatccac  | 126120 |
| ccacctgggc  | ctcccaaagt  | gctgggatta  | caggagttag  | ccaccgcgcc  | cagccagaag  | 126180 |
| accctacttt  | tctatttggc  | ttcccacatc  | tgactgctag  | catagagcct  | gctcccagag  | 126240 |
| tttcataatt  | aaaaaacat   | gaatgcttct  | gagggactct  | ccaagttag   | ggtcagggta  | 126300 |
| ggtgcaaaag  | gaatgatgtg  | acctgttgtg  | tttccctttt  | tcccttgact  | tccaggaagc  | 126360 |
| tctgccgttg  | ggtcactgca  | cagcccctgt  | cttttatgtg  | gcgtagccag  | ttagctcagt  | 126420 |
| cctgcgggtg  | agtccactag  | acttctagaa  | ggaacagact  | ggagcaggct  | cctcctcagg  | 126480 |
| ctccctccac  | tctccctggc  | tgccagtgcc  | catcttacca  | ttggcatgga  | agtccagggt  | 126540 |
| ggaggttgca  | tttccaaggg  | agttggtggc  | tgagcacttg  | tactccccac  | tgctcagttt  | 126600 |
| ctccaggtct  | cggatctcca  | ggcgcagggg  | gttgggacca  | gaggtaccac  | tgaagcgtgg  | 126660 |
| gctgtgatca  | ctgtccccgg  | aggtggaggc  | caggatatga  | cccccatgtg  | acagcaccag  | 126720 |
| tgtggccagg  | ggctcactga  | ccacagagca  | gtgaaggatg  | cccacaagtc  | ccgcctgggt  | 126780 |
| ctccaggaag  | gctgtcaggg  | ctggagttag  | aggcgggtct  | gcgtggagac  | gagaggtggg  | 126840 |
| cctgtcacc   | tcagacaagg  | gcattccctg  | gataccctga  | tacaaccctg  | gacctctgca  | 126900 |
| ccgctttgtc  | ccacactgcc  | ctgcgagaag  | ggggtgatcc  | caaagtgcct  | gagtgcctgc  | 126960 |
| tgaatacact  | tttgtccttg  | gctgggctgg  | gtacgtcact  | ctgttgtccc  | agtcagtgtt  | 127020 |
| caaggccacc  | ctgcaagtgg  | ggataaaaagc | cccacttcaa  | agatgagaaa  | actaatagag  | 127080 |
| agacctggtg  | agggacagca  | cctcagccag  | gtgaccaaag  | tgagcatcat  | cagcaatggg  | 127140 |
| acaagctgac  | acagagtgcc  | tccgtacagg  | gcagagcagc  | atgtccgcgg  | ccttcccacc  | 127200 |
| cagagggcat  | ggtctcaggc  | cattcaagtc  | aggcaaagt   | ggagtgaggg  | acctcctcac  | 127260 |
| aacagcaggc  | ctgactctg   | caagtttcaa  | cgtcaagaat  | gacaaggaaa  | gactgaggaa  | 127320 |
| cagttctcaga | ctaacggaga  | atgaagagac  | gcaacgaccc  | aatgcaatat  | gtgaactgtg  | 127380 |
| attgtattct  | ggaccagaaa  | aaaaatggct  | acaaaagaca  | gtattaggtc  | cactgggtaaa | 127440 |
| atgtgaatat  | agattatagc  | ttagataaca  | gtcttctatc  | agccagggtg  | agtggctcat  | 127500 |
| acctgtaatc  | ccagtacttt  | gggaggccga  | ggtagggtgga | tcacgtgagg  | tcaggagttc  | 127560 |
| aagaccatcc  | tggccaacat  | ggtgaaaccc  | tgtctctaca  | aaaacacaaa  | aattagccag  | 127620 |
| gcaggatggt  | gggtgcctgt  | aatcccagct  | actcaggagg  | ctgaggcagg  | aggagaatcg  | 127680 |
| cttgagcccc  | agaggcggac  | attgcagtga  | gccaaagatag | caccattgca  | ctccagcctg  | 127740 |
| ggcaacacag  | tgagactcca  | tctcaaaaaa  | aaaaaaaagt  | cttctatcaa  | tgctaatttt  | 127800 |
| cctgatcatg  | atcattgcat  | tgtggatatg  | taagagaatg  | atcttaggaa  | tttagcggta  | 127860 |
| aagaggcctc  | atgtctgcaa  | cttcaggaat  | ataaatatgt  | aggtagatac  | ataaataaca  | 127920 |
| tatgcatatg  | tatatagagt  | gtccatatat  | gtataagtgc  | acatgtccac  | atagagtggg  | 127980 |
| tgtgcataca  | cacaagtgca  | cctgtatata  | tgcaagtcta  | tatccataca  | tttatatgta  | 128040 |
| tgtatgtgog  | tgtgtgtgtc  | tgtgtagaca  | gcgagaaaga  | ttaagcaa    | gtgccaaaat  | 128100 |
| gttaataaat  | gggaaatcta  | ggttaaaggt  | atataatcat  | tatttgtctt  | tctctgcaac  | 128160 |
| ttttcataag  | tttaaacctc  | caaataataa  | ataggaggga  | actaggggagg | tcaaagcatt  | 128220 |
| gcccagggtg  | gcagagctgg  | gacatgagct  | caaggccacc  | tccaggtaag  | tggccttgaa  | 128280 |
| gttcccatgc  | ccaggacccc  | aaccttctct  | tggggctcca  | ccatctggtc  | ctgctctgac  | 128340 |
| tgtgtccagt  | gcccacccac  | cctggcctct  | tacggttgac  | taccaogctg  | acagggcccg  | 128400 |
| agcgctcgct  | gccaacggag  | ttctgcacct  | cacagaagta  | gaagccagta  | tcagccctag  | 128460 |
| tggccaagtg  | cagccggagg  | gtatgggagt  | gggcatcctc  | cagcaggaca  | tggttcttgt  | 128520 |
| accagctgta  | gcggagatca  | ctgggtgcct  | cggtgggtgt  | ggtgcagact  | agtgtcactg  | 128580 |
| tctggttctc  | caggatggga  | cctgctgggc  | tcacctggac  | ctcagccact  | gcaagggcag  | 128640 |
| catagggagt  | gctggggggg  | cccagccaac  | ttccagcccc  | agccccatac  | agtccccggg  | 128700 |
| tctcagccaa  | tcagcctcaa  | tctctgcaca  | tcctacacca  | cccgcctgct  | gctacagggg  | 128760 |
| agcctcctgg  | agtgtctctg  | atctctttgt  | cctgtctacc  | aggatcccc   | tgacacccca  | 128820 |
| cccttcgtgg  | gggtattgtc  | acctgtctca  | tggttaaggca | gggctgggac  | tccacacctg  | 128880 |
| catccaggaa  | gcactgggtt  | atgccagtgt  | ctggccctgc  | cctgcccctgt | ctcccctccg  | 128940 |
| tccttgaggc  | ctgagctcct  | ccctattctc  | tggccaatac  | gcaaccttcc  | caagaactca  | 129000 |
| ctgaagatgt  | ggaggctgat  | ggggggtgag  | accaaagagc  | ccagccggtt  | ctcagcttgg  | 129060 |
| cagggtgtaga | cgccagcatc  | gctccaggct  | gcctggggca  | ggtgcagcac  | accagtcctg  | 129120 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| gtttggaggc  | gtaccccatc  | cttgagccac  | ttaatggaac  | tgactgcagg  | gtagctgctg  | 129180 |
| ttcacctggc  | aggtgagtg   | gaccagctca  | cctggaagga  | tgttccctcc  | cgaggggctg  | 129240 |
| aggaggatct  | tcacacccct  | gggggcatct  | gcaagtcaca  | gtagggggta  | ttgggtaaag  | 129300 |
| tgcttgggga  | gggcagagga  | tggcacactt  | cttcttgccc  | ccctttaaaa  | gctcagtcct  | 129360 |
| aaggaagtat  | gcccagataa  | aacagcagtc  | cccttcaatc  | ctcaccaggg  | gcatgctctg  | 129420 |
| tctgcccctg  | tccgtccctt  | tcccgcctcc  | tgcgcctgat  | gcattcctat  | gcccattgag  | 129480 |
| agctgatcat  | gtgacgcttg  | gccagcgtcc  | agcctacccc  | accagctgta  | gttttctgct  | 129540 |
| tcccaatgct  | gtccattgct  | cctgctctat  | ttgggatgag  | ctccacacac  | agggtcacgg  | 129600 |
| gtaccccact  | gctagcttta  | gtggcctgtg  | gaagctattg  | gtaagggacc  | aactacctag  | 129660 |
| tgggaggggg  | ccaaaggcag  | catcaaacta  | gctctgaaaa  | tagttaccca  | gtttgtaagc  | 129720 |
| aagaggccaa  | caacacaaag  | aactgcattt  | ccttaaatct  | tgttccaaag  | cctctctctg  | 129780 |
| tgtattcaag  | tgttttatte  | ttatTTTTTT  | agaaagagga  | tctggctcag  | tcagccaaag  | 129840 |
| tggagtgcag  | aggcacaaatc | atagctcact  | gcagcctgaa  | actcctgggc  | tcaagtgatc  | 129900 |
| ctcctgcctc  | agcctcccga  | ggaactggga  | ctacaggtgc  | aagccaccac  | atccagctaa  | 129960 |
| tttttgtttt  | ttaatttttt  | tgtagagaca  | gggtctcact  | atgttgccca  | ggctggctctc | 130020 |
| gaactccttg  | cctcaagcaa  | ttctcctgcc  | ttggcctccc  | aaagcactgg  | ggttacaggt  | 130080 |
| gtgagccact  | gtatccggcc  | tcaagtgttt  | aatttgtgcc  | aggcactctt  | ctaaatcctt  | 130140 |
| gacctgggtc  | atctccttta  | tttgtttttg  | ttgttgttgt  | tgttttttgt  | ttgagacaca  | 130200 |
| gtctcactct  | gtcaccaggg  | ctggagtgc   | gtggcacaa   | cacaactcaa  | tacaactcca  | 130260 |
| cccccggggt  | tcaagcaatc  | ctagtgcctc  | agcctcctta  | gtaactggga  | ttacaggcat  | 130320 |
| atgccaccac  | accggctaa   | ttttgtattt  | tttagtagag  | acgaggtttc  | accatgttgg  | 130380 |
| ccaagctggg  | cttgaatccc  | tggcctcaag  | tgatccaccc  | gcctcagcct  | cccaaatgta  | 130440 |
| tgggattaca  | ggcataagcc  | accgcgccc   | gcccattctc  | tttaattttt  | atcataaatc  | 130500 |
| tgtgagacag  | gaaccatcta  | ttgttatctt  | cgtcatagac  | tgagaaaaca  | gagccaggca  | 130560 |
| ggaaagggat  | aaatcccacc  | tcccagcat   | cctccagcac  | ggggtaaata  | ccggggcact  | 130620 |
| tgctgccaac  | cctgaagctg  | catgggagct  | cctgttcccc  | aaccctttgt  | gtctcccttt  | 130680 |
| ctctgcccag  | cctggcctca  | aggacaagcc  | ccttcgaagc  | agtaggaggt  | tgagggaacc  | 130740 |
| atttaacgaa  | gtccttgggg  | ctcagcagtg  | tctctccact  | tgcctacccc  | tggaggccca  | 130800 |
| ggagcagcct  | gggttttgcat | caggagcaag  | ggttccgttt  | ctgtgggctg  | gagaggggct  | 130860 |
| ggtttctgtc  | aggagcaaca  | gatgcgctca  | gccacaagg   | tgtgtcccca  | gacaactcac  | 130920 |
| acttcacttg  | agggtgaatc  | tcgtctctgag | ccctgtgatt  | ggccacggag  | agctggcagc  | 130980 |
| gcaggatccg  | gccgtgggtc  | tgccaggaca  | tggccatgtg  | gaggggtctc  | aggtggccga  | 131040 |
| cgccgggtgg  | ctcaaacttc  | tggctgttga  | aggtgacaga  | gcgagcaggg  | tcctggcctt  | 131100 |
| gccactgcag  | tctgacctgc  | tcctgcaggg  | atacgtaggg  | agtggagcag  | ttgaagtcca  | 131160 |
| cctctgtgcc  | ctcgagaagc  | tccaccgggg  | aggcaatgg   | gggcaccctg  | ggctcctctg  | 131220 |
| aggacagaga  | cagcagtgct  | caggaccgc   | ttttgccacc  | cctgagatcc  | ctcgccctgg  | 131280 |
| aaaccccagc  | tgaggagaga  | gccttgggga  | gggggctttt  | aggggaagga  | aagacgcttt  | 131340 |
| ctactccagc  | cccacttggg  | tgaatacaag  | ggagaaccag  | gcctgggcct  | acgccggggc  | 131400 |
| tggaacagag  | gctgagactg  | gctgggggta  | gattcaggac  | aagggtctgg  | gctgagagcc  | 131460 |
| aaggggtcca  | gaagcagctt  | gggaatccct  | ccgggggggc  | agccaggcca  | cccacttat   | 131520 |
| cacctgttac  | tgtgaaccaag | gtgcctttca  | catctgacca  | gcgggttgacc | tactgatct   | 131580 |
| cgaagcggaa  | gttgtaggaa  | ccagagtcc   | cgggctgcag  | gtccttcagc  | agcaggttgc  | 131640 |
| acaccctgtg  | ctcgggggtc  | cccatgaact  | cgggtcggcc  | gcggaagcgg  | gcctccacca  | 131700 |
| gcttgggggtc | cgccgagtg   | ctcaccacct  | gccgctggcc  | cgagtagtcg  | tagtaccaga  | 131760 |
| tggcctgtgat | gccgtcgggc  | acctccacgt  | cggcagggaa  | gctgaagatg  | caggggataa  | 131820 |
| gcaggcaaga  | ccccttcaca  | ccctgcacgt  | cctggggact  | ggagaagccc  | catgaggcct  | 131880 |
| ggcctggggg  | aagaacggca  | gggggacaga  | ggggaggggtg | atacaggcct  | cagggtgcca  | 131940 |
| cagagccctg  | cgacctgccc  | cagagaaggt  | gccccagctg  | ggctcccaaa  | ttctgccctg  | 132000 |
| ccccgagaaa  | tgcacactta  | gagcagccct  | tctcagtgcc  | ccaggggtca  | gtccactgcc  | 132060 |
| cgaggctgct  | cagagggttg  | gtagggtggc  | tcgaagacag  | atggcagctt  | cctgccacgc  | 132120 |
| tcttggccac  | tgcctgattg  | ggcctccct   | tgacctcagc  | aaccaaaccat | gtgaccacag  | 132180 |
| catagtctaa  | tctgccaaag  | gctggactta  | tgaaggctgg  | accctggaag  | acaggcacta  | 132240 |
| ggcccgcaaa  | gcttacctgc  | tgggaagaat  | gaggccagga  | ggagaagctt  | gggcaagaag  | 132300 |
| cccatagcag  | gttcttctgtc | tgctcctgtt  | gcctaagagg  | gtggtgcgca  | ctgcgctggc  | 132360 |
| tgggctcaca  | ggggcctcca  | gggacacctc  | tgggcacttt  | agccccagca  | cctgctagaa  | 132420 |
| gtccgagcct  | gtgtccccac  | ctcctctgct  | ggccaaccca  | ataagagggc  | agggctctta  | 132480 |
| aagacctctg  | agtccagacac | cagcagagag  | ccaggaggcc  | acgttcccag  | ctcaggctgt  | 132540 |
| gcccaggaat  | gcccctcactt | ggtggcctgc  | ctcagaaaag  | cctgtgtgtc  | ccttggccct  | 132600 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| atcccaagtt  | ctgctttccc  | agccccctcaa | ggataccacc  | ctaaggcaga  | tgaacagctg  | 132660 |
| ttttctccctc | ttccccactt  | ccctgcccc   | tccccaccac  | ccaaaccaac  | aggaactgga  | 132720 |
| gcccagagat  | gcccagttac  | tactccaag   | acaccagct   | agaatgatgg  | tttcttcctg  | 132780 |
| aggcttgtct  | cctaccacct  | gccttactaa  | ctatagacca  | taatggggct  | ttactgaact  | 132840 |
| tgcgaagtg   | ctgcctttaa  | cagtcactcc  | cctgctcaaa  | aaccttctgt  | ggctcccat   | 132900 |
| tgcctgtgag  | atgtgaaaag  | tcctcatttc  | ctgccccag   | ctctgtcccc  | attccctgct  | 132960 |
| ctccgcagac  | ccactctggg  | ggcagttctg  | tctgctcaag  | ggctccctag  | ctgcccagct  | 133020 |
| ctatctccac  | cacagataat  | ctttgcctgc  | tgaacctta   | ttcaaccttc  | aaggagcagc  | 133080 |
| atgaatttgg  | cttcacagctg | gaacttctcc  | tttgagggtc  | ctgtagctac  | cccagagcta  | 133140 |
| tctctatttt  | ccttgccctg  | ttttttacag  | cttgtagagag | cccatttctc  | aggacctaga  | 133200 |
| actgaagata  | tgtgccccat  | agcagtgcgg  | agcctaccag  | gcattcagca  | aacctcttag  | 133260 |
| tgactaagag  | aggggtgagg  | tctttagggg  | ttcagagctg  | aggttcagag  | ttggagtggg  | 133320 |
| gaggtggcaa  | ggcaagtcta  | ggtttgaaag  | gtagcatgag  | agcgtgtgg   | aacacatacc  | 133380 |
| ccacaaatat  | gagctcaatg  | tgtgcggagt  | gtaccatata  | caaaaaggca  | ggccctcaac  | 133440 |
| catggagtgc  | ccctgggtcag | ggagtgtcta  | aggggtacca  | tagacctgag  | cccaaaagga  | 133500 |
| agagatgcc   | gaaacacata  | taagtgaaac  | tataaaactc  | ttagaagaaa  | acaggtgaaa  | 133560 |
| atcttcatga  | ccttggatta  | ggcaatgctt  | tcttaagtat  | gaaattaaaa  | gcacgaaaaa  | 133620 |
| aaaaaatagg  | taagttggac  | ttcatcaaaa  | tttaaaactt  | ggccagacac  | agtggctcat  | 133680 |
| gcctgtgaac  | ccaacacttt  | gggaggctga  | ggcaggagga  | tcacttttagc | ccaggagttc  | 133740 |
| aagacgagcc  | tgggcaatac  | tgcaagactc  | tgtctctacg  | aaaaatttaa  | aaacaggcct  | 133800 |
| ttggatccag  | ctactctgga  | ggctgaggtg  | ggaagatcgc  | ttgagcccag  | gggaggggctc | 133860 |
| gaggtgcag   | tgagccatga  | ttgcaccact  | gcactccagc  | ctgggtgaca  | gagcaagacc  | 133920 |
| ctgtctgaaa  | agaacaaaca  | acagctgggt  | gcgggtggctc | acgcctgtaa  | tccagcact   | 133980 |
| ttgggaggct  | gaggcatgca  | gatcacaagg  | tcaagagatt  | gagaccatcc  | tggccaacat  | 134040 |
| ggtgaaaccc  | cgtctctact  | aaaaatacaa  | aattagctgg  | gtgtggtggt  | gtgcacctct  | 134100 |
| agtcttagct  | actcgggagg  | ctgaggcgag  | agaatcacct  | gaacccagga  | ggcggagggtc | 134160 |
| acgggtgagcc | aggatcacgc  | cactgtactc  | cagcctgggtg | acagagttag  | actcttctgt  | 134220 |
| ctcaaaaaaa  | aaaaaaaaaa  | aaaaggatat  | gaatcaaagg  | acactatcca  | gagagtgaat  | 134280 |
| aaaaggagga  | caaccacag   | aatgggagaa  | aattatttga  | aatcatttat  | ctgataaagg  | 134340 |
| actaatatcc  | agaatatata  | aagaattcct  | acaatgaaca  | acaacaacca  | ccaaaaaaac  | 134400 |
| catgaaatcc  | aactccaaaa  | atgggcaaaa  | cacttgaata  | aacatttctt  | caaagaagat  | 134460 |
| atataactgg  | ctgataagga  | catgaaaaga  | tgctcaacat  | cactaggcat  | taggaaatgc  | 134520 |
| aaatcaaaac  | caaaccacag  | tgagatacca  | cttcacatcc  | gttagaatgg  | ctattcacaa  | 134580 |
| acaaacaaag  | caacacagaa  | aacaataaat  | attggtgagg  | atgcgaagtt  | gaaattcttg  | 134640 |
| tgtattgctg  | gtgggaatat  | aaaatgggtc  | cgtcactgtg  | gaaaacaatt  | gggtcattcc  | 134700 |
| tcaaaaagtc  | aacataggat  | taccatatga  | tccagcaatt  | ccactcctag  | gtatataccc  | 134760 |
| aaaagtactg  | aaaacaggga  | ctctaacaga  | gtacaccaat  | gttcacggca  | gcactattcc  | 134820 |
| actaaaaggt  | ggaaacaggt  | caagtgtcca  | tcagtgaatg  | gatgtggata  | aacaaactgt  | 134880 |
| ggatgtgaca  | tacaatggaa  | tatcaatcgg  | ccataaggag  | gaatgaattc  | taacatatgt  | 134940 |
| gaaccttgaa  | aacattatgc  | tcagtgaaac  | cagccagaca  | caaaaaggga  | aatattgtag  | 135000 |
| ggttccaatt  | acatgaaata  | tctagactac  | gtatattcag  | agactgaaag  | tagaatagga  | 135060 |
| tagaggtaac  | caggggctgc  | agggaggggg  | agctaattgt  | taatgattgc  | tgagtctctc  | 135120 |
| agataatgaa  | aaagttctgg  | aaatagtgg   | gatgggtaca  | caatattgca  | aatgtacct   | 135180 |
| atgtcatggg  | gtgtacactt  | aaagacagtt  | aaaacagtaa  | attttacatt  | atgtatatatt | 135240 |
| caccacatac  | acacccatgt  | tgccaacttt  | gcaatcctcc  | ctggtcctaa  | atgctgactt  | 135300 |
| ggccaagtga  | acgaggaggc  | tggaaagtggg | gacaggaaac  | tcatgacctc  | ccagctccca  | 135360 |
| gcccacccgc  | ctcaggggct  | gggctcagca  | gattccaatg  | actaccaggg  | gtcacacctg  | 135420 |
| ggaagggggg  | gagccgaggc  | ccaggggcag  | tcagggtgac  | caggtgggac  | ttagcctgct  | 135480 |
| gcagaaggca  | gaaggtgccc  | cagcaggggg  | cacagtacag  | ggcgggattg  | ggacaggaag  | 135540 |
| gacaccgctc  | cccagggggc  | ccagccctct  | cgcaggctgc  | tggagtggac  | tgatctggcc  | 135600 |
| atztatggag  | gcccaggggc  | tcactctccag | ttctctagga  | agccctaggc  | ctcctcctct  | 135660 |
| tctgggaaga  | tgcaccccca  | gcctccacac  | caggttcttg  | gccactggag  | aatgatatag  | 135720 |
| ctggggccct  | gggacctgga  | cacctcaccg  | tgaagaaaag  | cagcctgctg  | ggcacactgg  | 135780 |
| gggtcagatg  | tgtccctggc  | cacaggggat  | gtcagggcca  | gctctgctat  | ggccaggggca | 135840 |
| gctatcttgt  | cccagctccc  | ctgttccctcc | catttggggg  | cctgaaaagg  | gcaatcgtga  | 135900 |
| acctgatgga  | agaatgggtg  | ggttctggac  | acagcgaccc  | tggaaacagg  | cgcgggggag  | 135960 |
| gaccttttcc  | aggaccactc  | ccatcacata  | atgtagaggc  | cacctatgct  | tagcccggcc  | 136020 |
| ctaaccocaa  | aggggtcagc  | cccaccggaa  | tccagcctat  | tggctcagcc  | tgtcaccaca  | 136080 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aagggcagct  | tcagcccaga  | taactgttct  | ggaaacagaa  | agagcagggg  | ccgctcagaa  | 136140 |
| aggagatctc  | tgtccctgtt  | tgaaagcctg  | gagttgaggg  | gacagtgcce  | cgccccccgc  | 136200 |
| cccccgcaac  | ttgggttgca  | gctgtggcct  | agttagcacg  | cagcgcccc   | tggtgggtcga | 136260 |
| gggggaattg  | cgggtcccgg  | gaaagggggc  | ggtgtgccag  | caacagggag  | caggcagctc  | 136320 |
| tgcagccctg  | aaccatccct  | cccttgggtg  | actcttttgg  | aaatcattgt  | ccccagaca   | 136380 |
| ggaggttcc   | gaggttcata  | cttgggtctc  | caagtcttgg  | gtgctctgaa  | gacaggattt  | 136440 |
| taaatcccca  | ctcctactat  | tggtatgtgt  | tgcatacagg  | tggcttgagc  | tggcctggta  | 136500 |
| cacagtgggc  | actcacgttt  | gctgcctgtt  | tgagaccaag  | tgcctcagga  | ggtcttggag  | 136560 |
| ggttggctgg  | ggccccaagt  | ccctgacctc  | tgattccaga  | ggccaagttt  | agctggggaa  | 136620 |
| gaaagggcag  | aggcagtttc  | cctatggaca  | gctaggcccc  | ggtgtaggat  | tcagtttctg  | 136680 |
| tttccctgaca | ccaaggcttc  | tccccaaact  | cccacttggg  | ctagagaagg  | aagaacacag  | 136740 |
| ggtgacatgg  | ccagctggag  | gttactggcc  | cacagatagg  | gagtcagggt  | acggatgggc  | 136800 |
| aattcctgga  | gcagattatg  | gtcaaaatag  | tggaaatccc  | caatcacagg  | ccaaatgttt  | 136860 |
| aattctcaga  | gccatagaat  | ccataactaa  | tgcttattgg  | ctttgacatg  | ggcagtagaa  | 136920 |
| atttcacact  | tcattcctta  | atctggctct  | aaatgcttct  | ggctggagtg  | ctcactctcc  | 136980 |
| aaactgtgct  | ggacagcacc  | agaagccctc  | ctgtggacgg  | acgaagtggg  | gatggatgaa  | 137040 |
| gtggaattgt  | gctgggggta  | gagtaaggag  | agattatcgt  | tgggctgggc  | tggggctgag  | 137100 |
| gtcaggggtga | tgatcaggtt  | ggaattctgg  | aagcaaatta  | aggctgggat  | tgggggtggac | 137160 |
| attgaggttg  | agtgatgggt  | gggggttaagg | gtgaaggttg  | gagttgggtg  | caggtgatgg  | 137220 |
| ttaagatacg  | gtggaggctg  | ggttgagatg  | aggatgatag  | agtcggagtt  | gtgggttaggg | 137280 |
| ttgggatgat  | ggatgggttg  | gattagatga  | gatgagtaaa  | tgggttaggg  | tgggggtgcaa | 137340 |
| gtgaggtgag  | ggtgagataa  | ggttggaata  | ggggctgggg  | ctggggctga  | ggttagggtca | 137400 |
| gagcaggggtg | atggtcggga  | ctgggattgg  | gatgcagggt  | ggaaggattt  | gggggtggtg  | 137460 |
| gaaggggttc  | agttgagtc   | atctttgatt  | agtgtcttgg  | acttgggttg  | gtttgggggtc | 137520 |
| caccactcgc  | accagatgg   | agcccccccc  | cgaccctctg  | ccctatcccg  | ctcagccagt  | 137580 |
| ttcagcccag  | ccccgctcct  | aatgctccac  | tcaccctctg  | ggcccaggga  | ccaggggacag | 137640 |
| gggtacctgc  | tccacagaag  | gaagtggctg  | cggcgggtgct | ggacctgccc  | agaggagaa   | 137700 |
| aggaaggacg  | gccaagagct  | cctggtgcag  | ctggctcccc  | agggtctctg  | cggtcaaga   | 137760 |
| gagaaggatc  | ccgtatcagg  | ggctgcttcc  | tctttcccaa  | agcctcagct  | ctactgtcca  | 137820 |
| accagaggc   | tggctcaggga | ggcagctgca  | ggccttgcgc  | aatgccaaaa  | cggggaagac  | 137880 |
| ctccataggg  | gaaggccctc  | ggcaaggcca  | gggacttagg  | gactccagca  | agcagaagtg  | 137940 |
| ggacctctgc  | aacgctggag  | cctccccagg  | caaagtgaga  | aatggagtgg  | gggactccca  | 138000 |
| ttcaccagcc  | aaatccacac  | cccactctct  | ctgagccctt  | tagggagggt  | gggggagggtg | 138060 |
| gaagggggct  | tcctgcacac  | agctctctct  | cctgacacct  | gagagggagg  | cgcgcccagc  | 138120 |
| ctgggggtggg | gatgcactac  | acgatgcccc  | gaccaaggca  | gttattctcc  | aagccattca  | 138180 |
| ggagcctccc  | tgtaccactg  | agtactctta  | agaaccccca  | agaggcagtc  | ccgtgttgtg  | 138240 |
| ggggacataa  | ggccccaatg  | tccaggacac  | tctcagggtt  | ttctctttcc  | cctctcactg  | 138300 |
| tctgtacgt   | tttttggttt  | ttggcttctt  | gttgttgggt  | tttttctttt  | tgttttgttt  | 138360 |
| ttgcttttga  | gatgggtctc  | cactctgtcg  | tccaggctgg  | agtgcgggtg  | cacaatctca  | 138420 |
| gctcacggca  | acctccacct  | cccagtttca  | agtgatcctc  | ttgcctcagc  | ctctcgagta  | 138480 |
| gctgggatta  | caggcatgca  | ccaccacatc  | cggctaattt  | ttgtattttt  | ggtagagacg  | 138540 |
| gggtttcacc  | atgttgtcca  | ggttgggtct  | gaactcctga  | ccccaaagta  | tctgcctgcc  | 138600 |
| tcggcctccc  | aaagtctctg  | gattacagag  | gtgagccgct  | gcacccggcc  | actatcctat  | 138660 |
| accttcaccc  | ccaccttggg  | acaggaaagg  | aaagccccc   | ccacagctag  | ttactgttac  | 138720 |
| attactgcag  | cccatcttat  | tgagggccta  | gtttgtgcag  | gcacttgaca  | tgtgaagcag  | 138780 |
| atgttaattg  | ctccaagcaa  | tcccagatac  | aagcaacaag  | actttttttg  | gtttttgttt  | 138840 |
| gtttgtttgt  | ttttgttttt  | tgagacagtg  | tcttgcctct  | tgcgccaggc  | tgggggtgtcc | 138900 |
| tggtgcaatc  | ttggctcact  | ttggctcact  | gcagcttcga  | aattctaggc  | ttaagcaatc  | 138960 |
| ctcctgcttc  | agcctcctaa  | gaagccggga  | ttacaggcgc  | ttgccactat  | gcaoggctaa  | 139020 |
| ttttttaagt  | tattttgtag  | agacggagtc  | tccctatgtt  | cccaggctgg  | tctcaaacac  | 139080 |
| ctgggctcaa  | gtgactctcc  | cacctcgacc  | tcccaaagtg  | ctggaaactac | aggcgtaagc  | 139140 |
| caccactcct  | ggccattttt  | ttttttaatt  | tcacaactca  | aacttaattc  | aactcagtc   | 139200 |
| ctgcctacct  | gtactgggtg  | ggagctggcc  | acagcaaatt  | aggccctgtg  | acctgaggg   | 139260 |
| caaggccacg  | tgtgcagggtg | taaaaggatg  | tgaaacccta  | aatcagggtg  | gatccagaat  | 139320 |
| cgcaggccat  | ggtgccccaa  | agcagatgtc  | tggtgacatt  | ccaccctgaa  | atgctcaggc  | 139380 |
| tacagagata  | ttaggtctct  | atcactctgt  | tctcttttat  | agctcctgtg  | tccactccta  | 139440 |
| tgcttggggc  | atttctctct  | tgggcaaaaa  | cacaaagggt  | tcateccatt  | acttctctcc  | 139500 |
| tcaacagctg  | gtccggagac  | accagctct   | aggcctgtgg  | ggttgtgaca  | catgggtacc  | 139560 |



|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aatccttcag  | tccactggga  | ctctatacat  | ccacccttg   | gcttcattggt | gggaaagcat  | 139620 |
| cccttgactt  | tgaccttagt  | catatgactt  | gctctggcca  | atggatgtgc  | acggacatga  | 139680 |
| cacaaactaa  | gtcttgaaac  | gtacttaagc  | agtttgctt   | tccccctgaa  | tttctgtcat  | 139740 |
| tgccatgcaa  | ggggcaggct  | ccaactaacc  | tgctgggtcca | aagaggatga  | cgaacacatg  | 139800 |
| cagatgacct  | gaatcagacc  | catggattga  | aacaaaactc  | agctgagccc  | agcctacgtc  | 139860 |
| caccagacca  | gtcaacctgt  | ggatgcgtga  | atattattgct | ggatgctgct  | gagaattttg  | 139920 |
| tggctacatt  | agcaagatga  | tacaaggcct  | aagtcccaga  | acaacacacc  | cagaacttgc  | 139980 |
| ttacctttcc  | ttagcatgag  | gagagcaaag  | acttgtctac  | cttgattagt  | cagggagcac  | 140040 |
| tgcttctctgt | catttcccttg | agtatacagc  | aaactaggta  | aataaataaa  | aataactagg  | 140100 |
| taggctgggc  | acggtggctc  | acgcctgtaa  | tctcagcact  | ttaggaggcc  | gaggtgggca  | 140160 |
| gatcgcttga  | ggccaggagt  | tcaagaccag  | cctgggtcaac | gtggcgaaac  | cctgtctcta  | 140220 |
| cgaaaaatac  | aaaaattagc  | tgggcctggt  | ggcaggcgcc  | tgtagtctca  | gctactcagg  | 140280 |
| aggctgaggg  | acgagaatcg  | attgaacccg  | ggaggtggag  | attgcagtga  | gccgagatca  | 140340 |
| cactactgca  | ctccagcctg  | gatgacagag  | cgagactctg  | cctcaaaata  | ttttaaaaaa  | 140400 |
| tgtaattttct | cagtaggtca  | cactgtttaca | cattcatcta  | ataacaatta  | ttctttttct  | 140460 |
| tttttttttt  | tttttttttg  | agatagggtc  | tcactgtcac  | ccaggctgga  | caggctggag  | 140520 |
| tgcaatggca  | caatctcggc  | tcactgcaac  | tttgacctcc  | taggctcaag  | tgatcctcct  | 140580 |
| gcctcagcct  | cccaagttgc  | tgggactaca  | ggtgagtacc  | accagacgca  | gccaattttt  | 140640 |
| gtattttttt  | gtagagatgg  | ggcttcacca  | tgttgcccag  | gctgggtctca | cactcctgag  | 140700 |
| agttccaggt  | caagtgatcc  | acccaccttg  | gcctcccaaa  | gtgctgggat  | tacaggtatg  | 140760 |
| agccaccata  | cccagtgaat  | tattctcgtt  | ccagatagga  | aaaccaagtc  | ataggggaggt | 140820 |
| tagagaattt  | gccaaagaca  | aaactttttg  | gttgaaaaaa  | aataagtttt  | gtacaagta   | 140880 |
| tagaaaacac  | caaataacgg  | tgttttaaat  | aaaatagaag  | tgtttcaccc  | tctccctcaa  | 140940 |
| gtaagtgttg  | gcatccaaga  | tgatatgaca  | actccacaat  | catgaaacta  | gatccctttt  | 141000 |
| tatttttttg  | ctcagtcatt  | gtcaatgggc  | tgcttccagt  | cattgtccaa  | agtggctgct  | 141060 |
| tgcgctccag  | ccactgtatc  | tgcattcggg  | aaagcaggat  | ggaggaaagg  | acaaagtagg  | 141120 |
| ccatgcccct  | tactttaaga  | taactcctta  | aattgctgat  | gtgggtgggtc | actcctgcaa  | 141180 |
| taccagccac  | tcaggaggct  | gaagcaggag  | gttcacttga  | acccaggagc  | ttgagggtct  | 141240 |
| agtgaaactat | aattgtgtca  | ctgcattcca  | gcctgagtga  | cagagtgaga  | tcttctctct  | 141300 |
| taacaacaac  | aaaaaaggta  | attcttttaa  | ttgtacactt  | tattgctgct  | tatatccac   | 141360 |
| tagaagctag  | aagttagtca  | catggtctta  | agtctttatt  | ctaacagggtg | atgtgcccga  | 141420 |
| gtgaatcctg  | ggggccaggt  | gctaagaagt  | agaggggagat | gaagccctgc  | cctcctgtgg  | 141480 |
| agcaaccagg  | accttaattc  | agtcattcat  | tccttggacc  | taccagatt   | ccagcctcag  | 141540 |
| gcccctgccc  | acctccttgc  | cagtatctct  | gcagagcctc  | ctgcctcccc  | cagaggggtc  | 141600 |
| cagagccctc  | tgcttccctca | gccttcagac  | ccacttgetg  | tgtctcaggt  | ggggacaact  | 141660 |
| cagtctatag  | aggcccaagg  | gagatccttg  | agaccatac   | tgtagtccagc | ttggggcaga  | 141720 |
| tgggagtcac  | gtagctcact  | ggctaaaagc  | ctaggtcctg  | catagagctg  | ggttcaaata  | 141780 |
| caggcatatc  | catcacctgc  | tctacgatcc  | caggagtcac  | tcaatgactc  | agaagctaga  | 141840 |
| gtcctctgga  | aagcttgtat  | gaagagtcc   | cagggcctgg  | aacatcaaaa  | gagctctgag  | 141900 |
| aatgtttggct | ctgctgttgc  | ttctcttatt  | gggtatctga  | aggccgtact  | ctcctctctc  | 141960 |
| ctccacccga  | ggtgacctct  | actctcatag  | ccacaccctg  | gacctcact   | cattcagtca  | 142020 |
| cctcaccag   | tctgtcatct  | gtccctctgg  | ctctcctctc  | ctgttcatgc  | atcttgact   | 142080 |
| caccacttct  | tgtccctgag  | tgtttcaagg  | ctctgggggt  | ttcagagaca  | ctgaagggac  | 142140 |
| ctccctcctc  | aacacaacca  | caaggtctag  | gtggaatgac  | ccactaaggg  | accggctctg  | 142200 |
| aagccagaga  | cttcagggga  | aagtcaacaa  | gccaaggat   | gcccgttaca  | agaaagttaa  | 142260 |
| aagaccatg   | tacatgtcct  | cccgttttat  | tcctgtctca  | gggtctgggc  | atacagtgga  | 142320 |
| acacatgcag  | tccccaagg   | acaccgtctg  | tacagagtca  | gatggagtta  | agaacattta  | 142380 |
| gccggctggg  | cgcggtggct  | cacgcctgtg  | gtcccagcta  | ctcgaggagg  | tgaggcagga  | 142440 |
| gaatggcggtg | aacccgggag  | gcggagcttg  | cagtgagecg  | agatcgcgcc  | gctgcactcc  | 142500 |
| agcctggggc  | acagagacag  | actctgtctc  | aaaaaaaaaa  | aaaaagaata  | tttagtctgg  | 142560 |
| gcacggaagc  | ttatgcctgt  | aatcccagca  | ctttgggagg  | ctgaggtggg  | cggataacct  | 142620 |
| gaggtcagaa  | gtttgagacc  | agcctggcca  | acatggtgaa  | accccatctc  | tactaaaaat  | 142680 |
| acaaaaatta  | gccaggcatc  | tgtaatccca  | gctactcagg  | aggctgaggg  | aggaaaatca  | 142740 |
| cttgaaccca  | ggaggtggcg  | gtcacagtga  | gccaaagattg | tgccagtgca  | ctccagcctg  | 142800 |
| ggtgacagag  | caagactcca  | tctaaacaca  | cacacgcaca  | cgcacacata  | tttaaaccac  | 142860 |
| cacaccaaca  | tctagtccaa  | gatggtggac  | tgagaacttg  | tctctgccat  | tcctggccca  | 142920 |
| tccaatacca  | ctgagagcac  | agtaagcaaa  | gggaaaagga  | gacagaaggg  | ctgggaacag  | 142980 |
| gatggctggg  | ggatgggaag  | tatccactgc  | acaggatttt  | gatttaattc  | tagaagatag  | 143040 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| aaagaggggag | gatcacgttc  | aggaacagat  | gcgggttaaag | gaaaccagag  | ccaaagcacg  | 143100 |
| ctgagagaaa  | gctgccccag  | aggccggagc  | agaagtggac  | tctctacagg  | gactcaatac  | 143160 |
| accccaaaagg | gttggttagct | ggcacacgta  | cctctccacc  | cccacatgta  | acactgcagg  | 143220 |
| gcagaggaaa  | taccctaggt  | gagttccagt  | atcaatcaac  | ctgccctttg  | ttcataaaga  | 143280 |
| tgaattgggt  | accaggtatt  | accagacagg  | tgaggaagac  | caacacaaaag | agaaagatcc  | 143340 |
| agaaacaaaac | aggccagggc  | cattgggtca  | cgcctgtaat  | cccagcactc  | tgaggaggcca | 143400 |
| agatgggtgg  | atcacctgag  | gtcaggagtt  | caagaccagc  | cttgccaaca  | tggtgaaacc  | 143460 |
| ctgtctctac  | taaaaataca  | aaaattagct  | gggtgtgggt  | gggttctcct  | gcctcagcct  | 143520 |
| cccgagtagc  | tgggaggctg  | aagtaggagg  | ttcacttgaa  | cccaggagct  | cgaggctgca  | 143580 |
| ttgaactatg  | attatgtcac  | tacattccag  | cctgagtgac  | agagtggat   | cttgtctctt  | 143640 |
| aacaacaaca  | aaaaaggtaa  | ttctttaaat  | tgcacactac  | attgctgctt  | atattccact  | 143700 |
| ggctagagggt | tagtcacatg  | gtctcatgtc  | tttattctaa  | cagggtgatgt | gcccagagtaa | 143760 |
| atcctggggc  | cccagtgcta  | agaagtaggg  | ggagatgaag  | ccctgccctc  | ctgcttgaac  | 143820 |
| cctagaggta  | gagggtgcag  | agagcggaga  | tcatgccact  | gcactccagc  | ctgggtgaca  | 143880 |
| gagtgagact  | ccatctcaaa  | aaataaaaaa  | taaataaata  | aataaataaa  | taaataaata  | 143940 |
| aataaaaaacc | tacagcaaag  | aacaaagagc  | tattgcaccg  | ttactgtggg  | cctggcagta  | 144000 |
| ttatgacaac  | tttttttttt  | tttttttttt  | ttggagatgg  | agtatttttc  | tgtcaccagg  | 144060 |
| gctggagtgc  | agtggctcaa  | tctcagctca  | ctgcaacctc  | tgctcctggg  | gttcaagcga  | 144120 |
| ttctcctgct  | tcagactccc  | gagtacctgg  | tattatagge  | acatcccacc  | acacccggct  | 144180 |
| aatttttgtg  | tttttagtag  | agaccgggtt  | acaacatggg  | ttcaccatgt  | tggccagctc  | 144240 |
| agtatcaaac  | tcctgacctc  | aggtgatcca  | cccgccctcag | cctcccaaaag | tgccaggatt  | 144300 |
| acaggcatga  | gccaccacac  | ccatccagtg  | ttctgacaat  | tttatactta  | ttaactcatt  | 144360 |
| tagcaacctt  | ctgagctaga  | tgctattgtt  | atctccactt  | acaggaaact  | gagtcacaga  | 144420 |
| atgggtacagt | aaaataaacct | tgaccgaggt  | ttacatggct  | ggtaagtggc  | agagaaatga  | 144480 |
| aacttgaacc  | caggcaacct  | gtttccctgag | ttggactctg  | aactactcag  | ccatactgcc  | 144540 |
| tattttattta | tttatttatt  | tacttactta  | cttattttatt | gagatggggg  | cagccgggca  | 144600 |
| tggtggctta  | cgcctgtaat  | cccaccactt  | tgggaggctg  | aggggaacag  | atcacttgag  | 144660 |
| gccagcagtt  | cgagaccagc  | ctggacaaca  | aggagacagg  | gtcttgctct  | gtcaccagg   | 144720 |
| ctggagtgca  | atgggtcaat  | catggttcac  | ctgagcttcg  | acctcccagg  | ctcaagcgat  | 144780 |
| cctcccacct  | cagcctccca  | agtagctggg  | accataggca  | cccactatca  | tgtccagcta  | 144840 |
| atcgaaaaaa  | aaaaattata  | tagagatggg  | ggtctcacta  | tattgcctag  | actggtctct  | 144900 |
| aactcctggg  | cttaagcaat  | ccaccacact  | tgccctccca  | aagtactaga  | attacagggt  | 144960 |
| tgagccaccg  | tgctgacca   | atacttcctc  | tttaataaac  | taaacaaaaa  | gtgaacaaac  | 145020 |
| caatcccggg  | gggaacagat  | aattcaagga  | atacaagaaa  | acttatgaaa  | agaaagattc  | 145080 |
| tagtgccaat  | atggtcataa  | aacaaaaaaa  | agcagctatg  | aaaagaaagc  | aaaaaataat  | 145140 |
| tcttggaat   | ttaaaataga  | attgctgaaa  | taaaaagaaa  | ttcaatagaa  | aagggtgtga  | 145200 |
| acatctgtgg  | ttttgggttc  | caaaattagt  | tcactcttct  | ttgggtaaaa  | ataccctgat  | 145260 |
| tttctgtctg  | tattctcaag  | ccctgtgggt  | tgggtggaatt | gacacctctt  | ccatttactg  | 145320 |
| cacgagagaa  | gcatgttgct  | gctcacagtt  | cattatgagg  | agccagcagg  | taaagacaac  | 145380 |
| actgagggca  | gatgaaagat  | atatactgac  | cctgggtcct  | tgtggacatc  | attgagtccc  | 145440 |
| tggatcaagc  | cttacctgaa  | gctaaaggga  | tgagtgtctc  | tcactgttaa  | tagccactga  | 145500 |
| tcaaattcca  | tagaaggact  | ccaattatcc  | ttgcctcagt  | tgtgtgtcca  | gccctcggat  | 145560 |
| gagctaccat  | caaggggaaac | tgccagacct  | tggtgacaag  | cccacccact  | gaaaatgggg  | 145620 |
| aaaaaaccca  | agtcacatca  | ttgaactgaa  | agagctgcag  | ttccccgaaa  | gacaggagag  | 145680 |
| aagggaatgct | ggataaacac  | aacaactact  | tcatctcact  | gcacctgcaa  | agaatccagg  | 145740 |
| cacaggaagg  | tgggcacagc  | aagtagtctt  | tctgggtgac  | ataacaaaga  | caacatctgt  | 145800 |
| ttcttctctc  | agaaagaagt  | aagaaaggca  | attggagaac  | cagaaacgac  | aaaatctatt  | 145860 |
| ccaggaccaa  | atagaaagaa  | aacacatctg  | aagcatgatt  | ttgaacaatt  | agtgggggtgt | 145920 |
| aagaaaagag  | aatccatttg  | accttgacat  | ctgaaacagc  | aatatttctc  | catcaaggca  | 145980 |
| cttttagggta | gaggaaatga  | tactatgttt  | ccttctcaaa  | ggagctatct  | ggttacgtgg  | 146040 |
| ttctgcagta  | aatgacgttt  | atacaatcat  | aacatccgtt  | ggttttcaat  | tttcagaatc  | 146100 |
| aacctgaaga  | taaagcatgg  | aagatttcat  | tataattacc  | gaacagcatg  | caaatgttac  | 146160 |
| aaactttgac  | catgaaaatg  | taaaaaagag  | cccaggcgca  | gtgggtcacg  | cctgtaatcc  | 146220 |
| tagcactttc  | ggaggccgag  | gcgggtggat  | cacaaggcca  | ggagtttgag  | accagcttgg  | 146280 |
| ccaactagtg  | gaaaccccat  | ctctactaaa  | aatacaaaac  | ttagctaggc  | atgggtggcac | 146340 |
| gcacctgttg  | tcccagctac  | atgggaggct  | gaggcaggag  | aatcacttga  | accaggagg   | 146400 |
| cggagtttgt  | gggtgagccaa | gatcacggcg  | cggcactaca  | gcctgggtcaa | cagagcaaga  | 146460 |
| ctccatctca  | aaaaaaagaa  | aagaaaagaa  | aagaaaagta  | gacagcagaa  | attagagggga | 146520 |

|             |             |             |            |             |             |        |
|-------------|-------------|-------------|------------|-------------|-------------|--------|
| gatggcaggt  | gacctggggc  | tgggggaggg  | acaatttttc | ttacatagtg  | atagggtcca  | 146580 |
| agaaattctt  | tctaaaattg  | atgtatgagg  | aacgaaattt | taaatataga  | ttgttttagaa | 146640 |
| ctataagtag  | caccaacaca  | ggactccagt  | aatcacacag | caaacaaaac  | tgagaaaaca  | 146700 |
| gacatgggcg  | gggagatggg  | aatgggtgat  | ttccattccc | tgctttaatg  | atgaggtgtc  | 146760 |
| agtagatgct  | ctcttgagtt  | actaaattga  | gaaacaaaga | taaagtatat  | tgtttagagg  | 146820 |
| agtgatgttc  | aaactccaga  | aaaggctggg  | cgcggtggct | catacctgta  | atcctggcac  | 146880 |
| tttgggaggg  | cgaggtgggc  | agatcactta  | agcccaggag | tttgagacca  | gcctaggcaa  | 146940 |
| catgatgaag  | ccctatctct  | actaaaaaaa  | aaaaaatac  | aaaaaattag  | caaagcatgg  | 147000 |
| tggtgcacac  | ctgtagtccc  | agttattcag  | gatctcacta | ctgtactcca  | gcctgggtaa  | 147060 |
| tgagagtggg  | accctctctc  | aaaaaaaaaa  | aaaaaaaaaa | aaaagaagaa  | gaagaagaaa  | 147120 |
| actccagaat  | atataaaaaa  | gaaaagaggg  | cggaagagtg | gaactgcggg  | gtagaaaatt  | 147180 |
| ttgcattttg  | tttcatctct  | ctgtatatgt  | tgaacttttt | ttttaaccta  | catgtactac  | 147240 |
| gtcatatcca  | cagcccccaa  | ccatccacca  | ggcatcaact | gctgtgttta  | tatggagggt  | 147300 |
| tgagcaatca  | ttcctgcctc  | cttccagttc  | gtcttctgt  | actgcagagg  | ctagaaaact  | 147360 |
| aaattttatat | caccagatt   | cccttcagc   | taccttaatg | ctaccacctc  | attcagccat  | 147420 |
| cattgattct  | ccacttcacc  | aaactggtca  | ataattgtct | cttctaaaaa  | tattagaatt  | 147480 |
| gggactaaga  | gacactacag  | ctgaccctta  | aacaacatgg | gtttgaactg  | tgctagtctg  | 147540 |
| cttatacaca  | gattttttaa  | aaaataaaca  | tattgaaaaa | aattttggag  | agatgtgaca  | 147600 |
| atttgaaaaa  | actcacaac   | cacatagcta  | gaaatatcaa | aaaaaaaaaa  | actgaaacag  | 147660 |
| agttaggtag  | gtcaggaatg  | cataaatata  | tatgttaatt | ggctgtttat  | attattagta  | 147720 |
| gggcttccag  | tcaacagaag  | gctattagta  | attaagtatt | tgagtcaaaa  | gttatacaca  | 147780 |
| gattttttgat | gatgaggagg  | attagtgtcc  | ccaaccccc  | tggtgttcaa  | gggttagcta  | 147840 |
| tatttgaaat  | ctgctggtag  | ctggaattgg  | gcaactaaat | tcacgagatg  | tggaacgggtc | 147900 |
| gtatacatgg  | agaataaata  | aataaataaa  | taaataaata | agtatttgtt  | ggcctgggtgc | 147960 |
| agtggtcat   | acctgtaatc  | ccaacacttt  | gggaggccaa | ggcaggcgaga | tgacttgagg  | 148020 |
| tcaggagttc  | gagaccagcc  | tggccaacat  | ggtgaaatac | cgtctctact  | aaaaatacaa  | 148080 |
| aaatcagctg  | ggcgtggtag  | tgtgcacctg  | taatcccagc | tgcttggggg  | gctgagggat  | 148140 |
| gagaattgtc  | tgaacctggg  | agtcagaggt  | tgcaagtgcg | cgagattgca  | ccactgcact  | 148200 |
| cagcctggg   | gaaagagtg   | ggactctgtc  | tcaacaacaa | caacaaaaaa  | gcaattgtta  | 148260 |
| aaagaataag  | aaaacaagcc  | acagattggg  | agaatatatt | tgcaaaaaca  | atatctgatg  | 148320 |
| acccaaaata  | cacaaagaac  | tcttaaaacc  | caacactaag | aatcaaaaca  | acccatttaa  | 148380 |
| aaaatggaca  | aacgggcccag | gcatggtggc  | ggtggctcac | atctgtaatg  | ccagcacttt  | 148440 |
| gggaggccga  | gcagggcaga  | tcaccttagg  | tcaggagttc | tcaactagcc  | tggccaacat  | 148500 |
| agtgaaccg   | tctctactga  | aaatacaaaa  | attagccgac | cgtggtgacg  | tgcatctgtg  | 148560 |
| gtcccagcta  | cttgggtgtc  | tgcggcagga  | aaatagcttg | aaccaggagg  | gttgcaatga  | 148620 |
| gctaagatcg  | cgccactgca  | ctccagcctg  | ggcaacacag | tgagactccg  | tctcaaaaaa  | 148680 |
| aaaaattggg  | caaaagatgt  | gaacagacac  | ctcatcaaag | aagatatata  | gatggaagat  | 148740 |
| aagcatatga  | aaatatgctt  | aacctgatgt  | cattaaggaa | ttacaaattg  | aagcaacaag  | 148800 |
| gtaccactaa  | acccctatta  | aaatggtcaa  | aatccagaat | gctaaaaaca  | ccaaatgctg  | 148860 |
| gtgaggatgt  | ggagtgcacg  | gactctcatt  | cgttgctggg | ggaaatgtaa  | aatggtatgg  | 148920 |
| ccatttttaa  | gacagtttgg  | cagtttctta  | caaaactaaa | catagtctta  | ccatacaatc  | 148980 |
| taacaattac  | actcctagat  | agttacccaa  | ttgaggtgaa | aacttatatc  | cagggccgag  | 149040 |
| cacagtggct  | cacacctgta  | atcccatcac  | tttgggaggg | caaagaggaa  | ggattgcttg  | 149100 |
| aggccaggag  | ttctttcttt  | ttattttatt  | attttattat | tttcttttga  | gacagagttt  | 149160 |
| cgctctgtca  | cccaggctgg  | agtgcagtgg  | agtgatctca | tctcactgca  | atctctgctt  | 149220 |
| ccgggttcca  | agcgattctc  | ctgtctcggc  | ctctgagtag | ctgctgggat  | tacaggtgca  | 149280 |
| cgcaccatg   | cccagctaat  | ttttatatatt | ttagtagaga | cagggtttca  | ccatgtcagc  | 149340 |
| tagactggto  | ttgaactcct  | gacctcaagt  | gacctgcctg | cctcgggtctc | ccaaagtgtc  | 149400 |
| gggattacag  | gtgtgagcta  | ctgtgctcag  | ccaaggccag | gagttccaga  | ccagtccctgg | 149460 |
| caacacagtg  | atactctgtc  | tctacaaaaa  | aaaatttttt | taattagtca  | catgtggtgg  | 149520 |
| cacacacttg  | tagtgtcagc  | tactggggag  | gctgaggctg | aggatcactt  | gagccccagg  | 149580 |
| agtttgagg   | tgcagtgagc  | cacgattatg  | cctctgcact | ccagcgtggg  | caacagagta  | 149640 |
| agagaaagaa  | agagagagag  | agagagagag  | agaggaagga | aggaagggaag | gagaaaagaa  | 149700 |
| acgaaaagaa  | aagaagggaag | gaagggaagg  | aaggaggagg | aggaaggagg  | gaagggaagga | 149760 |
| agaagaagaa  | aaaggaaaga  | aataaaaaaa  | gaaaagaaaa | gagagagaga  | aaccttatat  | 149820 |
| ccatacaaaa  | ccctgcgcaa  | gaatgtttat  | agctgcttta | ttcataattg  | ccaaaaactg  | 149880 |
| gaagcaacta  | agaagctctc  | caatgggtga  | aaggataaac | aaactgtggt  | atatctgtgc  | 149940 |
| aatggagtat  | cattaagtaa  | taaaaagaag  | acttagcaaa | ccacaaaaag  | taatattgcat | 150000 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| actaaatatz  | catttagtaa  | tattaatatg  | cattaagcaa  | taaaaaaaga  | cgtgtcaagc  | 150060 |
| cacaaaaagt  | aatatccata  | ttactaaatg  | aaagaagcca  | gtctgaaaag  | gctacgtgct  | 150120 |
| gtgtgatttc  | aactatttga  | ctttctggaa  | aaggcaaaac  | tacagacagt  | aaaaagatca  | 150180 |
| aggggttgccg | gggggttctgg | ggacaggaaa  | ggatgaatag  | gtggagcacg  | ggaatttttag | 150240 |
| gtcagtgata  | ttattctata  | tgatactata  | atgggtggatc | catgatatgc  | ttttgtcaaa  | 150300 |
| acccatagaa  | agtacaacac  | aaagagtgat  | tcttaatgta  | aactatgtgc  | tttagtcaac  | 150360 |
| aatgtattga  | tattagttca  | ttaatttttaa | caatgtacca  | cactaatcca  | agatgttaat  | 150420 |
| aatggggaaa  | ctgcggtgtga | gggaggggat  | acatgggaac  | tctgtactac  | agctcaataa  | 150480 |
| ttttataaat  | ctaaaaacttt | tttttttaaat | aaagtctctt  | agaaaacaat  | aaaaataaaa  | 150540 |
| taaaaagtca  | atttttttttt | tttttttttga | gagggaaatct | tgctctgtca  | cccaggctgg  | 150600 |
| agtgcagtg   | cacaatctct  | gctcactgca  | acctctgcct  | ccttagttca  | agcgattctc  | 150660 |
| ctgcctctac  | ctcccagggt  | caatcaattc  | tcccacctca  | gcctcccaag  | tagctgggac  | 150720 |
| tacaggcatg  | cgccaccacg  | cctggctaata | gtttgtattt  | ttagtagaga  | tggggtttcg  | 150780 |
| ccatgttgcc  | caggctgggtc | aagaactcct  | gacctcaggt  | gatccgccc   | cctcagcctt  | 150840 |
| tcaaagtgt   | gggattacag  | ggttgaacca  | catgcctggt  | ctaaaatgtc  | cattttttaa  | 150900 |
| aggcagctctg | gctggggcag  | tggctcacgc  | ctgtaatccc  | agcaccatgg  | gagaaggctg  | 150960 |
| agcagggtgt  | atctcttgag  | cccaggaggt  | cgaggctgta  | gtgtgctatg  | atgggtgccac | 151020 |
| tgactccag   | cctgggtgac  | agagtgaagc  | tctgtctcta  | aaataaataa  | ataaaaaatac | 151080 |
| aataaaattt  | tataaaggcaa | tctgcagcaa  | gagaagatga  | agttagcagg  | agagcaggac  | 151140 |
| agaggagagt  | ccatgaacaa  | caagtggtca  | aaagtggtca  | tgggtggcagc | tcccaggtc   | 151200 |
| ctggcatatg  | ttcaattccc  | tgttcccagc  | cctcaggaag  | cccagatgtc  | cccacctgcc  | 151260 |
| cctacatata  | aacctgggat  | ccttgacatc  | aacttctcct  | ttccttcatg  | tgtctaatg   | 151320 |
| agtttttgtt  | acttgtggaa  | catttaacta  | acatcattaa  | ccaagtggac  | ctgtctcaga  | 151380 |
| gtgggtgttat | gagaagacag  | cagccaaaag  | acagctgcag  | ccaagcacag  | tggctcatgc  | 151440 |
| ctgtaatcct  | agcatttttg  | gaggccgagg  | tgggtggatc  | acctgaggtc  | aggagtccga  | 151500 |
| gaccagcatg  | gccaacatgg  | cgaaaccccc  | tctccactaa  | aaatataaaa  | attatccggg  | 151560 |
| tgtgggtggcg | agcgccctata | atcccagcta  | cttgggaggt  | tgaggcagga  | gaattgcctg  | 151620 |
| agcccagggg  | gcagaggcgg  | cagtgcagcc  | ggatcgtgcc  | acttcactcc  | agcctgggtg  | 151680 |
| aaagagcaaa  | actctgtctc  | aaaaaaaaaaa | aaaaaaaaga  | cagctgcaac  | aatgtcaag   | 151740 |
| ttctgtgtgt  | tttcttttct  | tttctttttt  | ttctatttaa  | ttaattttatt | ttagagtcat  | 151800 |
| agcctcccta  | tgtcaccacg  | gctggagtgc  | agtggcacag  | tcacagctca  | ctgtagcctc  | 151860 |
| aacctcctgg  | gctcaggcga  | tccttccacc  | tcagcctcct  | tcctagctgg  | gaactacaggt | 151920 |
| gtgtgccacg  | acatctggct  | tgtgtgtttt  | cttttctttt  | tttttgagac  | ggagtcttgc  | 151980 |
| tctgccaccc  | aggctggagt  | gcagtggcgc  | gatcctggct  | cactgcaacc  | tctgcctcct  | 152040 |
| gggctcaagc  | aattctcctg  | cctccgcctc  | ctgagtagct  | gggaatacag  | gcgcacacca  | 152100 |
| ccatgcccag  | ctaatttttg  | tattttttagt | agagacgggg  | tttcaccatg  | ttggccagga  | 152160 |
| tggctctcat  | ctcctgacct  | tgtgatccac  | ctgactcggc  | atcccaaagt  | gctgggatta  | 152220 |
| caggcgtgag  | ccactgcacc  | ctgcctggct  | tgcatgtttt  | ctgacatact  | gtcaaaagga  | 152280 |
| tactcatact  | aaatggcaac  | acattctcaa  | gccccttccct | tttcttctcc  | tatctgcttt  | 152340 |
| accacacacc  | atagctgctt  | ttaccgtttt  | ctccttaaaa  | actcaaaaaa  | accttcccc   | 152400 |
| aacctatcta  | tccacttctt  | gtcctccctt  | caaaccctct  | ccattttgat  | tctgcatcat  | 152460 |
| aaaaaccaag  | cagagtctctg | gggaggcaga  | agcctcctct  | tatgatattg  | gaaggggggtg | 152520 |
| gatgaaattg  | agcacacaaa  | gccagaatcc  | tcttttgtgg  | aaagatgggg  | gcagtgggca  | 152580 |
| gaggaagca   | ggctcatctc  | tttccctctc  | ttcccttccc  | ttctctttca  | caccacacgc  | 152640 |
| tccgcctggg  | tgagctcatc  | gtcctcgtgt  | cttcaatttc  | cacctccagg  | agatcgactg  | 152700 |
| ccaaatttct  | accogcaatc  | ccaaactcca  | ctctgagctc  | cagacccatc  | tactaattgc  | 152760 |
| ctagtcaaca  | ttttctctgt  | gggatcgaa   | cagataggat  | tatattctgc  | tgtgactaac  | 152820 |
| agagactgaa  | aattcagagt  | cctaaaacaaa | agactttctt  | tttcttgtgc  | tgaaaagtct  | 152880 |
| ggaagtacac  | agccctaggc  | tggaaatagg  | atactgctgg  | ggggctccgc  | ctccttgcat  | 152940 |
| gctgcccctc  | tgtatctctc  | agggtgtgtc  | cctcaaccac  | actgtccaaa  | gtgggtgactg | 153000 |
| aaggaccagt  | cctccactctg | acattccagg  | caacagaatg  | gagaagaaca  | ggctgaggaa  | 153060 |
| ggaggaacca  | caggcgccct  | ccagctgttt  | caaggaaggt  | tcctggaaga  | tgctaaataa  | 153120 |
| caattccggt  | tatatctcat  | tagccaggct  | tagtcacgtg  | gcaacaccta  | gaagcaagaa  | 153180 |
| aggctgggaa  | atatactgtt  | tctgctgggt  | taccgtatgc  | atgactaaaa  | agagagcgct  | 153240 |
| ctgttcttag  | gaaaaatgaga | atagactggg  | ttgggggagg  | ataaccagca  | ttggctacac  | 153300 |
| tgcagcccca  | caacctcctc  | atgcttgatg  | tggctgaaat  | aaaaagagca  | tctcccccta  | 153360 |
| acatttctct  | tccttagtgc  | ttgttccctt  | ctgttaccoc  | ataccctagc  | agagggggcaa | 153420 |
| ttctgtgggg  | tgggctcact  | ggcaaggggga | ggtctcacc   | aagctcagtc  | aagggtgagc  | 153480 |

51/154

|             |             |             |            |             |             |        |
|-------------|-------------|-------------|------------|-------------|-------------|--------|
| agagagaatt  | gtacaaggaa  | gctgaatttc  | cagctggtct | ggaggaaaca  | ctgaaaaccc  | 153540 |
| caaacagaag  | atgtccaggt  | ggagcaaatg  | cagctgagag | ttcttgcca   | aggggacaga  | 153600 |
| atgtcagcag  | agtaggacac  | aagggcaggt  | ctctactgaa | agcaccaggg  | gcaaagtcac  | 153660 |
| ccagcccttc  | aaccggctgg  | ccaccaagca  | catcggtccc | actctgcctc  | ctgcccgctc  | 153720 |
| caccctttgc  | acttttgtct  | gtaccatctc  | ctccaccagg | agtgcctctc  | ccgctcctcc  | 153780 |
| atgtggtata  | ctctcagcca  | cccctcaggc  | tcagatccag | agccacctcc  | tccaagaagt  | 153840 |
| ctccacctgt  | ctgctgggaa  | ccctcacatc  | acactgggag | acgaagccaa  | aagtgggggg  | 153900 |
| gtcagtacag  | aggtgtggga  | ggtaggtcca  | atcctggctg | gacctggctt  | tggtgtgcat  | 153960 |
| ttgggacagg  | gtattgggga  | ggttcagcag  | gaaggtggga | cctcaaggct  | ttctttttat  | 154020 |
| gaaggaggag  | gaaccagctg  | ttgacagagc  | agcaacagca | gggagccacc  | aaggctgagc  | 154080 |
| cttgaaccct  | gggccaccag  | gctcctgcag  | gcaactggga | gcagctcagc  | ctctcacagt  | 154140 |
| cccattttcca | gggaggaaga  | tgatgcagcc  | tgaggcaatt | ggccacctgg  | aggaagtcca  | 154200 |
| gctgggtggg  | gcaggcagag  | gaattcctgg  | aaggaagctc | acccctcagc  | actggggtag  | 154260 |
| agggacacta  | gagacagggg  | tggcccagga  | gctgccacc  | tcctatcccc  | atgaataagc  | 154320 |
| aggtggggcc  | gggtgaccag  | ctgggagcca  | cccctgggtg | ctcaggtacc  | atgcttctgg  | 154380 |
| cccagcctcc  | tgagctgtgg  | cacaggcaag  | agcaccggcc | cacggggtct  | gcatggggca  | 154440 |
| gtgacctgtt  | ctctctaagc  | ttcagctttc  | tcactctgtg | aaagaggggtg | aagagtcctt  | 154500 |
| ccccaccatg  | gttgtggtga  | gggtaactga  | agtgtcaggg | taggggaccc  | agtggggcac  | 154560 |
| cagcacacct  | gaggttcagt  | caacaaggcc  | cactggtgaa | gaaaacctcg  | cccaccctcc  | 154620 |
| cgtccctcct  | gaggcaggtg  | gtggaggcac  | atgccctctc | cattgtgtgt  | gcattgagtg  | 154680 |
| gtgggtgcat  | gcattgtgtg  | gcattgtgtg  | gtgtgtgcat | gtgtgtgcat  | gcattgtgtg  | 154740 |
| gcgaatgtgt  | gtgtgtgctg  | acatgcatgt  | gtgctaagct | ctggagaaca  | agcagccagg  | 154800 |
| cttgggcagg  | agtccggggg  | caaggggggg  | acatgcaaag | ttctctccac  | agcccctcag  | 154860 |
| tgctaacagg  | ctgaggtggg  | ggatgacatg  | tcctagaaac | aaacttcacc  | ctctgcccta  | 154920 |
| aaactgcctg  | ctggtctgcc  | acatcggggc  | aagctgtggg | tacttggaag  | gagggccctg  | 154980 |
| ctcctgcctt  | ccaaacccag  | gaagtcttga  | aatccccatg | gaggggacct  | ggggccaggg  | 155040 |
| gactcccaa   | gcagacacag  | caccacacac  | aggttctaag | ggaccccttt  | ctccccaacc  | 155100 |
| gcctgaggta  | gggatggatg  | tggacaagca  | gctcagggtc | ggggctggcc  | taggtggaag  | 155160 |
| gttcagaagc  | acagaaggca  | ggacgtgccc  | aaggggccgg | acatgggggt  | ggaggctagg  | 155220 |
| ggctggaaaa  | agagtttggg  | aggtcagcca  | agggcagctc | acgccttggc  | tcctctctgg  | 155280 |
| gtcatgggga  | gccacggaga  | gtataggagt  | ggtgatggga | tggccccaga  | gttgtgcctt  | 155340 |
| tgaaagtcac  | ctctggttgc  | agatagagat  | ggactggggg | cagggtctgag | aggaagggca  | 155400 |
| tctcatgaga  | ggggaaatcc  | ctgaccacag  | tcagaggaa  | aacagcgggc  | ctggggagag  | 155460 |
| gctgctcctg  | gggaggacct  | actcagaggc  | tgattcttcc | tcgtgctacc  | ccaccagggg  | 155520 |
| agatgcagga  | agctggtgag  | gtcagtggtg  | gggccacaaa | ccatgctgga  | gaggagcagg  | 155580 |
| agggagacgg  | gagggcacag  | agaggttagag | ggcagggact | ctgaggacct  | tcattccatgc | 155640 |
| ccacctggac  | aattgaatcg  | agcttcctct  | agtgtggaat | gtgggtctac  | tggggtattt  | 155700 |
| tgctatcctt  | ctctcatgtc  | agactggatg  | cggcccaagg | gtgggtctctg | tgactccctc  | 155760 |
| atcagactga  | ggggggccga  | aatgtggggg  | tgggggaagg | agggagcttc  | caaacaggaa  | 155820 |
| ggtttgaggg  | gcgctggcac  | tgaggacagg  | ggatagcccc | cccatgagtc  | ttcaccctcg  | 155880 |
| ctcagagcct  | gctgccgtgg  | ccctcagaga  | cgagcccatc | tttgagagctg | gagttgggga  | 155940 |
| cagggaaacc  | aggtggaagg  | gaaaaagaaa  | aaaggggcag | gcatgagtg   | gaggggccag  | 156000 |
| accctgcaga  | cccaggaca   | tgcacagcag  | acacacactc | caatccctcc  | tggtctctgt  | 156060 |
| cagcctcaaa  | cgaggtgcag  | gcctgcaggc  | cccaaccaag | gacccctgtt  | acacacagac  | 156120 |
| acaaagacac  | aagatgcagc  | aacacacaca  | cgtaatacac | aaacacagtc  | acacactagc  | 156180 |
| gcacacacag  | acacagtcag  | acaatccaag  | aaacacactc | acacatgaac  | agacatataa  | 156240 |
| gcaaagtctt  | acacatacaa  | agacacactt  | aagcacagac | acaaacacac  | caacagatgc  | 156300 |
| acagacggag  | acacatagac  | gcacactggg  | aggcggcctg | tggagactgc  | tgattggata  | 156360 |
| ccatgttggc  | ttggtatcca  | gaattgttgg  | tgcagcgcac | tccccatgcc  | tgacctcatt  | 156420 |
| aaggaggccc  | tctggggcag  | tggtctctgag | aagaaaggca | gcaagcctag  | ggaggtgaca  | 156480 |
| ctgtagctgg  | ggagcatttt  | ggctggggcag | agaggaggaa | ggagagcatt  | tcaggggagg  | 156540 |
| aaaatagccc  | ttgcaaagg   | ttggagtggg  | agagcaaaag | ggatgtccca  | agaactgtgc  | 156600 |
| ctcatttgat  | ggggctgtgg  | ggcccttctc  | tgacgggcag | ccttcccagg  | cacaaaggga  | 156660 |
| ggtggactga  | gaagcatctg  | gaacttctgt  | gggggcacat | ggggcagagg  | ggaaggggaa  | 156720 |
| caggcagaac  | gagtcagctg  | cctttgaaag  | ctccacgagg | gacctagcag  | aaattcattc  | 156780 |
| attcatccgt  | tcatttcattt | atttatgaca  | ggttctctct | ctgtcaccce  | agctggagtg  | 156840 |
| cgggtggtgtg | atcacactta  | ctgcagcctt  | aaactcctgg | gctcaagcaa  | tcctccaccc  | 156900 |
| cagcctcctg  | agtacctagg  | actacagacg  | tgagctacta | cacctggcta  | atttttaaaa  | 156960 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| cttttttagag | acaggggtctc | cctttgttgc  | ccaggctggt  | ctttgttgcc  | cagccagctt  | 157020 |
| caagccatcc  | tcccacctca  | gactcccaa   | gcgtggatt   | acaggcatga  | gccacgggtt  | 157080 |
| cccagcccag  | aaagagacct  | tcaccccaca  | aacccctgt   | ctggaggcca  | agtcccagca  | 157140 |
| tcccagaag   | ggccccggtt  | aggacaagaa  | aaaggctgtg  | tgccctcggag | gagaggctgg  | 157200 |
| gtgccccctg  | gtgcggagt   | gtccccctgc  | cgggcatggg  | gcctcagtct  | gtgagactcc  | 157260 |
| gcgcctccgt  | cgtctgaact  | tgagcaaaga  | cagagctagg  | cctagcctgg  | ggaccccgagg | 157320 |
| gcctgccctc  | atccgcaatt  | ctgcctcttt  | tctggccact  | gaggaccggg  | ctagggaagc  | 157380 |
| tgtaggggca  | gtaggggtgg  | gaggagagag  | ggttccgggt  | caagggactg  | gaggggtctg  | 157440 |
| gggaggggaa  | gtggaggggt  | ctagggcggg  | cccttcccc   | tccgctgttg  | ggccctcgct  | 157500 |
| gaagggggccg | aggctgagcc  | ccaaggtggg  | cctgtggggg  | cggggcgggg  | gagccctca   | 157560 |
| gcagctggac  | cgcaggaggc  | cgaccagggt  | ctggaactcc  | tcttccccctc | cctccgcgcg  | 157620 |
| ccctctccc   | ttctcccacc  | ccctgggctg  | ctaggggagc  | gggggaggcg  | caggaggggc  | 157680 |
| tcagggagggg | gcccctggacg | cgggaccagg  | ctggggccct  | cggcgagggc  | ccgcgcaggc  | 157740 |
| aggccctgcc  | gctggcctcg  | cacatctggg  | ccgggagcgc  | gggcccagccc | ggcggcgcag  | 157800 |
| gcggcgcggc  | catccggcct  | gggggagggg  | gcggcgcgccg | gggtgggccc  | gccccaaagaa | 157860 |
| ggcggaaggc  | ctcagggccg  | ggcgcacgtt  | gagggctgcg  | accgcgcag   | cgggcacggc  | 157920 |
| ccaccgagct  | ggggggcccgg | gatcgcgcg   | gctggacggg  | gctggagctg  | tcgggagggc  | 157980 |
| ggaggtgagt  | tctggggcgg  | gggctgccgg  | gcgccccgag  | tggagaaagg  | cgaggaggtt  | 158040 |
| tgcctcccg   | ggctgtcggc  | tgagacccca  | ccaaaaacct  | ccagactttg  | tctggtgggg  | 158100 |
| acagatctgc  | aagccccctc  | ctgcagcgtg  | gctgcgggct  | cgggaaggca  | cttggggcag  | 158160 |
| cgaccttggt  | ctcccacctg  | ggcttcgggg  | gaccacctcc  | cctgtccctc  | gcacagctgc  | 158220 |
| tcagaggctt  | aggctggaga  | aactcctgtg  | tcttaggggc  | tggcaggact  | ggtgtgggt   | 158280 |
| ggggcctgct  | ccaggggagcc | aaactgaggg  | acccggtgcc  | taggtctgag  | gtggaaacct  | 158340 |
| gggagttgag  | tgtatgtgtg  | tgtgttgggg  | ggtgcctcag  | gggtcccagg  | ccctgtggtc  | 158400 |
| ccgaagcgca  | acagagtact  | gagtggccca  | ggtgccaggg  | tctgaggctg  | ggacctggtc  | 158460 |
| ttccggggcca | gcttgaggctc | caggtggggg  | ggggctccagg | ttcccgctag  | ggcagtggag  | 158520 |
| cctatctgtg  | aggttagcgt  | gtgttgatga  | gagagactgg  | aagggagtga  | tcacgaaact  | 158580 |
| agggaccctc  | cggagagcag  | acgcagcgca  | ggaatcccag  | gccagagggg  | taatgggggg  | 158640 |
| ctgtgggcag  | agcatgggag  | gggctctggt  | gtttgtgaac  | tggctgctcc  | catgatgagg  | 158700 |
| acccaagggg  | gtgcggcagg  | atgtatggaa  | agcaaagatg  | aggacaggga  | ggggccatgg  | 158760 |
| gaggccattt  | ctgtcactcc  | ctgccccagc  | gcaggatggg  | gcagccacct  | cactcagcca  | 158820 |
| cacatggtcc  | catcttctta  | agggctcctc  | tccagaacag  | agtggggcag  | ccctttgggt  | 158880 |
| gcggtttctg  | atctgtgctt  | tgggcccctg  | ttggcccagt  | ctgtggcttg  | gaagaaacgc  | 158940 |
| ctgtggcctc  | acttagcccg  | ctgacggagt  | cctgttccat  | tacagggctg  | gctgggtgct  | 159000 |
| gagcagagag  | cagcaggcgg  | aagaggcagg  | gtcctagtga  | ggaggggagc  | tggaatgacc  | 159060 |
| tgggtgggct  | ctttgtctac  | acaggcatcg  | tgtggagccc  | tgaaacccag  | aacactgcat  | 159120 |
| tcccatagt   | ctcctgggga  | catgagcgcc  | cagctgaaac  | acccacacac  | ggaagagatc  | 159180 |
| ctccgaggat  | tctgtcagge  | aaagggttgg  | gcccctgcaa  | ggtggctcac  | gcctgtaatc  | 159240 |
| ccagcacttt  | gggaggatga  | agtgggagga  | tcgcttgagg  | ccaggtggtc  | gaggctgcag  | 159300 |
| tgagccgtga  | tctgtccact  | gcactccagc  | ctaggcaaca  | gagcaagacc  | ctgtctcaaa  | 159360 |
| aaaaggagcg  | ggagtggggg  | cagggttcgt  | cctaggctat  | tgtaaagacca | ggggatactc  | 159420 |
| tgggatggga  | gtctctggtt  | gtgagttcct  | gctggggcac  | accctatcct  | ggtccccagt  | 159480 |
| gactaatcgg  | ccctgcctat  | tgggtagaag  | tcagggccag  | gggtcagggt  | gaagctgatg  | 159540 |
| actagaggct  | aatatgaagc  | caagagttag  | gggtcagtc   | atgggcaggg  | ttagggatca  | 159600 |
| ggctggggcc  | agagtttggc  | atgtgcttgt  | agccaggaaa  | tggagttoct  | ttctgtgtaa  | 159660 |
| atcaaaactgt | cccactcaat  | tactctcaaa  | cccataggca  | ccaccagtct  | gaaattagta  | 159720 |
| ggccccgtgaa | gcggggctca  | gcctccccctc | attttcagct  | gtggactctg  | cccagacctc  | 159780 |
| gctttcggga  | gctgggcttg  | tcccggtgtc  | tgtggtcagc  | actctcctta  | agtgtccctg  | 159840 |
| ggtgtgtgtg  | ctcgcttggc  | agtcagccac  | ctcttggtt   | acctccacct  | ctcggggaat  | 159900 |
| ccatgggatac | ctccttcttt  | aacacaattt  | tttttttttt  | tttttttggg  | gacagggtct  | 159960 |
| cactctgtca  | cccaggctgg  | agtgcagtgg  | tgtgatcatg  | acttaactgca | gcctcaacct  | 160020 |
| tccaggctca  | agcaattctc  | ccacctcagc  | gtccaaagt   | gctgggactc  | taggcagca   | 160080 |
| ctaccacgct  | gggctaatat  | ttgtattttt  | cgtagagatg  | gggttttgct  | gttggccagg  | 160140 |
| ctggtctoga  | actcctgagc  | tcaagcaatc  | caccactttt  | gacctcccaa  | agtgttagaa  | 160200 |
| ttacaggcat  | gagccaccac  | accagccct   | taacacaaat  | tgatttagtt  | cctacctcct  | 160260 |
| tccttaaaga  | actcatgatg  | actagaatgc  | gattctcagt  | aagacccaaa  | tatctaacaa  | 160320 |
| gaaaattcag  | aatctgcac   | atgaagggaa  | agaacactgc  | tcaaacctgc  | ccaactccct  | 160380 |
| gcttttatca  | ataccttgca  | tatctggact  | tgttaacata  | agccctctgt  | gtcaccatt   | 160440 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ctacaggaaa  | gggcactgag  | gcacagagat  | gtggcatgat  | gcactcacgc  | taggtgccat  | 160500 |
| aataggctct  | tgccatccct  | gaaaccccag  | cataggggca  | ctcagcatat  | cgcacaaatga | 160560 |
| gagacggtag  | ataactggac  | aagcttccga  | tagctgggaa  | gagtataggc  | ccatgggtttg | 160620 |
| gatgtggaat  | ccagtgtaga  | gcgaggactc  | agtctttttt  | tttttttttc  | ttgagatggg  | 160680 |
| gtctcactct  | tctcgcccag  | gctggagtgc  | aatggcgaga  | tcttggtctca | ctgcaacctc  | 160740 |
| tgtctcccag  | gttcaaacga  | ttctcctgcc  | tcagcctcct  | gagtagctgg  | aattacaggc  | 160800 |
| gtctgccact  | acgcttggct  | aattttttgta | tttttagtag  | agacaagggt  | tcaccatggt  | 160860 |
| ggccaggctg  | gtctcgaatt  | cctgacctca  | ggatgatccac | ctgcctcagc  | ctctgaaagt  | 160920 |
| gctgggatta  | caggagttag  | ccacggcgcc  | cggcccagag  | tctttttttt  | tctgggtgggt | 160980 |
| ctaggttggt  | tctcctccag  | gtacaatagc  | ctcggtccag  | aaaatccatt  | cattcattca  | 161040 |
| acaaagtggg  | tccccagtgc  | ttagcacaga  | gcctggccat  | aagaatctcg  | ataagtgttt  | 161100 |
| gctgaagaat  | cagaacaaca  | tacaaataag  | gctgccagct  | cctccctcga  | gaccttgatc  | 161160 |
| cagtagtatg  | ctgggttaatg | gtttccaatc  | agctctggaa  | gaaggataaa  | atcctgattt  | 161220 |
| atgggtgttg  | ccaattttctg | tcttgtaaat  | actcccagcc  | atgactgatt  | tgaagctacc  | 161280 |
| aatgtgatgt  | cattgaatgc  | agaattggga  | agagatagaa  | agaatcagct  | cgcagtagct  | 161340 |
| gggtgggacc  | agctcagcat  | cccactccat  | catcctgagc  | tgggctctgg  | gctcctgccc  | 161400 |
| ctggctccta  | agaacccctcc | cctccccctc  | tctgtgcttt  | ccctttttctc | ctcaacagcc  | 161460 |
| agctgctata  | tgcattgccc  | caaccagcac  | ccatccttct  | tctgattctg  | attcctgagc  | 161520 |
| atccaactcc  | aatggtagac  | aggacagcaa  | gagtgtctgg  | aaatctccat  | gggtgctggc  | 161580 |
| atgcaggaga  | atcctggaca  | tacaagagtg  | tgtgggagtc  | caggtgtctg  | tccacaattgg | 161640 |
| ggctgggttg  | tgtgaactgt  | gtgagcaagg  | gtgtgtgcat  | gtgtctatgg  | gcatgccaga  | 161700 |
| gccagctctg  | tgtctgcctg  | tgttctatag  | caaaactttc  | agattggggc  | gctggtgaca  | 161760 |
| gagaaggaat  | gcattggccc  | tgcgtagcct  | cctcctttgt  | ttagaggagc  | agccaggccc  | 161820 |
| accaatgaga  | acatagtgtc  | agtggctctgc | agactggctg  | agagacactg  | ttgagcaaca  | 161880 |
| cttctcaaac  | ttcaatgtgc  | acctgaatca  | cctgggaatc  | ttgttaaaat  | aggccagacg  | 161940 |
| cgggtggctta | cgcctgtaat  | cccagcactt  | tgggaggctg  | agacggggcg  | atcacctgag  | 162000 |
| gccaggagtt  | cgagaccagc  | ctggccaata  | tggtgaaacc  | ccgactctac  | taaaaataca  | 162060 |
| aaaattagcc  | aggcattttt  | aggtggcatt  | ttaggcaaat  | taggtggcgc  | atgcctgtgg  | 162120 |
| cccagctac   | tggggaggct  | gaggcaggat  | aatcacttga  | accagggaga  | tggagattgc  | 162180 |
| agtgaactga  | gaactgagat  | cacgccactg  | cactccagcc  | tggggggacag | agtgaactgc  | 162240 |
| tgcctcaaaa  | ataaataaat  | aaataaataa  | ataaataaat  | aaataaataa  | atatgaaata  | 162300 |
| catgttctga  | ttcagcaggt  | ctgggctggg  | gcctgaggtt  | ctgcatttct  | aacaagcttc  | 162360 |
| caggtgatac  | caatgctgct  | gatcctcagc  | ccacactttg  | agtaacaaag  | ctgttgatgg  | 162420 |
| tttgaaatat  | cctgcctata  | gcctggcccc  | agaacttcat  | agacaggatc  | gagccctctc  | 162480 |
| catcttagaa  | aataaatatt  | cttctctacc  | tcccttcagc  | tttagcgccc  | tccaaactact | 162540 |
| gcctcacaat  | gtattcctga  | tttcagccaa  | gctttttcaga | aattctagct  | caatatgatt  | 162600 |
| gcgtcactac  | actccagcct  | gggcaacaga  | gtgaaactcc  | atcttaagaa  | aagaagttct  | 162660 |
| agctcatgtc  | acctcttgct  | cactcctcac  | ctaaggcttac | ctcccttggg  | ccactaaaac  | 162720 |
| agtgtttcca  | gtctcatctg  | gggacctttt  | acagggatct  | agctttaacc  | atctcttggt  | 162780 |
| acaatcttca  | tgagaccatg  | aggcatctca  | gtctcctggg  | ccggtatata  | ttcctctggt  | 162840 |
| cagcctttaa  | atgtagggtc  | tctcctgac   | agagtgcac   | cctatctcaa  | aacaaacaaa  | 162900 |
| caacgaaggc  | caggcacggt  | ggctcatgcc  | tgtaatccca  | gcactttgag  | aggccgaggt  | 162960 |
| gggcggtatca | cctgaggtca  | ggagtctgag  | aacagcctgg  | ccaacatggc  | gaaaccccg   | 163020 |
| ctctactaaa  | aatacaaaaa  | ttagctgggc  | atggaggcac  | tcacctgtaa  | tcacagctac  | 163080 |
| tgaggaggct  | gaggtaggag  | aatggcttga  | acctgggagg  | cagaggttgc  | agtgaacca   | 163140 |
| gatcgtgcca  | ctgcattcca  | gcctgggcaa  | cagagtgaga  | ccctgtctca  | gaaaaagaa   | 163200 |
| aagacaaatt  | gagtcaatgcc | cccaccccac  | cctgccacct  | caaacctttc  | aagatttgct  | 163260 |
| gttgcccttc  | atataaaatc  | aaatctgaaa  | taggggtgcac | aggccctgcc  | tgcagggtaa  | 163320 |
| gccagaccca  | gccctctgtc  | cagacttacc  | atgtcctctt  | cccaccccg   | attatccaac  | 163380 |
| tttgaccttc  | ctagcactca  | gcaacgcttg  | tcattttcaa  | atcacttgcc  | tgattatctc  | 163440 |
| ctccatcaga  | cogttgggtt  | ctagggcagg  | gacagtgtct  | ctctgggtca  | ctatctccca  | 163500 |
| gcacctggca  | cagagcaggc  | accagaaaac  | atttagaagg  | atcctgattc  | caagaaagta  | 163560 |
| aagccatttt  | ttgacacaag  | ctgggaactg  | tgggacttta  | gatgagattt  | aaggaattac  | 163620 |
| tggttaatttc | attagacatg  | agaatgggtg  | tgcgatgaca  | ttgtttaaaa  | cgtctttatc  | 163680 |
| tgttggagag  | atgtattaaa  | gagttcacgg  | gtgaaacaac  | atgatggctg  | ggatttgctt  | 163740 |
| caaaggactc  | cagccaaaag  | taaacaaata  | aaataaataa  | gcattcagat  | gagggaacaa  | 163800 |
| agggacaat   | aaatcccact  | ctgcttgctc  | gttctctgtg  | acagctacag  | ggcctgcaag  | 163860 |
| gatgttggct  | gtccccggaga | tgggcctgca  | ggggctgtac  | atcggttaaga | acccccacca  | 163920 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ttctgccctg  | accatcacc   | cactcacc    | agcaggcact  | gagaccacc   | tacagggtgt  | 163980 |
| gcaatgtacc  | tctcccacag  | agcttcccaa  | cagctcttct  | gtgagtgagt  | gtgtgtgtgt  | 164040 |
| gtgtgcatgt  | gtgcagggca  | ggctgtgtgt  | tactggaccc  | ccctgtgtgt  | gcagaaaggg  | 164100 |
| tctgtgtgca  | ccggtctgtg  | tatgtgagta  | tgtgtgcaca  | caggactggg  | ctattcacag  | 164160 |
| gacctcatg   | tgtatatgag  | tgtgggctga  | gtgctgtgtg  | ttgtgtgtgc  | cagaaggggtc | 164220 |
| tgtgcacaca  | ggagtgtatg  | catgtgcgtg  | tgttttgttc  | aggctctgtc  | cacaggaccc  | 164280 |
| tgtgtgagtg  | tgtcagttag  | tgtgtgtctc  | tgcgggcagg  | tctctgtgca  | cggtatctgt  | 164340 |
| gtacgggagt  | atagtctgca  | gggtgggcatg | catacagaga  | gttctgctgt  | tgcattggagg | 164400 |
| gtctgtgtgc  | acaggagtgt  | ttgaatgtcc  | atgtgttttg  | ttcaggccat  | gtgcacagga  | 164460 |
| ccccgtgtga  | gtgcatcagt  | gagtgtgtgt  | tcttgcgggc  | aggctctgtg  | gcacaggatc  | 164520 |
| tgtgtacggg  | agtatatgtc  | gcagggtggg  | acgcattgcag | agagtctctg  | gtgtgcatgt  | 164580 |
| gcgtctggcc  | atagcctggg  | ctgggagcag  | gcttgggttc  | tccgcagggg  | tgggaggccg  | 164640 |
| gaggcaagag  | ggacactcct  | caggaagtcc  | attcacaag   | acctgggtgg  | ctgggcccct  | 164700 |
| ccttctctgc  | tcccttccca  | agctcgagcc  | acagcaactt  | ggcctctcgg  | gctgcagctg  | 164760 |
| gtggcagggc  | tggagcctgc  | cctggctgcc  | tgaggagagg  | gagcccttga  | aatgtggcca  | 164820 |
| ctttttccct  | gccccctccag | cacctcagac  | tctggccctg  | tccaggcgtc  | tccagagctc  | 164880 |
| ccctcatg    | ctggcccttc  | gctgtgttcc  | ccagggttacc | cccaaggcca  | gggagagtgg  | 164940 |
| agacctgcct  | ggcagcacgt  | gacaggactg  | tggctgaggg  | agcacagggc  | ctgcctggct  | 165000 |
| cagcctaagg  | cgagccagac  | ctgcctgatt  | tgtccattga  | gcttctacc   | agattctact  | 165060 |
| ctccctccct  | agccagcctc  | cctggcctgt  | ccatggatat  | caaggctgcc  | ccagtactcc  | 165120 |
| tcacatgccc  | cttagcaatg  | cccagtgtgg  | tgggtggcagc | taccagccag  | gcaggccctc  | 165180 |
| ctgcccata   | agccgcctga  | tctctggggc  | gaggcctggg  | gctgtcctcc  | agggtacttc  | 165240 |
| tactccctcc  | caccaaaggt  | ccagctcagt  | cccaggcgtg  | caagaggggtc | ttgatggggc  | 165300 |
| gtcctgcct   | gaaatgctgg  | ctcagggtgt  | gtgtatgtgt  | gtgtgtgggg  | gggtgggaatc | 165360 |
| agacttcta   | atgttgtgga  | cacccccagt  | tcaggactgg  | gctatctcta  | tgggtgtggg  | 165420 |
| aggaggtgt   | ccccctctc   | gatcctctc   | aagcagccag  | ctgggcctgt  | ctgtggccct  | 165480 |
| tgtgggcatg  | ggagctgcag  | agtctccctca | agtcctgatt  | agaaccacgt  | ggcccaactt  | 165540 |
| taggggtggg  | gggagcaacc  | ccacctccta  | gccagttagc  | tctgtagctc  | caccttttga  | 165600 |
| aggagtaggg  | cttctctctc  | ctctactccc  | tccaggagc   | tcacctgccc  | tatctccctc  | 165660 |
| gcccttccat  | ctctggactg  | agtctccctg  | ccacacaatg  | cttttctgtc  | tctctccctc  | 165720 |
| cccacactgt  | gtgtctgagg  | ggccctagtg  | gggagacagt  | gagctcctgg  | ggaaagcatg  | 165780 |
| ggctcatctgc | tagggaaacg  | ttgcctctga  | cagcagaccc  | tcagctcctg  | cccctgcgtc  | 165840 |
| atgtgtgtct  | ttctacacc   | ccgcagtcct  | cctccccagc  | agagagccct  | ttgccttagc  | 165900 |
| ccagcaagga  | ctgatggaga  | acctttgggt  | gcctggcttg  | gggccccgct  | gccataccta  | 165960 |
| ccctgcttgt  | gccaggatga  | actgcctctc  | cttctctgca  | ggctccagcc  | cggagcgggtc | 166020 |
| cccagtgcct  | agcccaaccg  | gctccccgag  | gaccagga    | agctgcggca  | ttgccccctc  | 166080 |
| cacacctcgc  | cagtctccag  | taagcccaga  | gcagggacca  | gggtgggtggg | gacctggctg  | 166140 |
| gtgtggacag  | ggctgtgcgt  | ggcaaaatca  | tgacagggcc  | tcagctagga  | aggaggaggg  | 166200 |
| gatgggggta  | gcacctgccc  | tcttgtcccc  | gtacactcca  | gtcatgtgcc  | cccagaaagg  | 166260 |
| attggccttg  | ggtgacctg   | gactataaag  | ggtaacaatt  | ctaactgagaa | ggcgagcccc  | 166320 |
| atgtaactaa  | ctcccaaccc  | ccccacacag  | gaggcaccca  | cggccattga  | agctggcctg  | 166380 |
| gctgaggcaa  | gtgttcccc   | gagcatgtgg  | cttttctagg  | tccccagagc  | actgggcctc  | 166440 |
| ttgtggacca  | cagcccacct  | taatccaagt  | catatctgaa  | ggagaggaga  | tgtgttgcca  | 166500 |
| aaaaaaaaa   | aaaatagtct  | ggaaaccatt  | agcttgggct  | gggcatgggtg | gctcatacct  | 166560 |
| gtaatcccag  | cactttggga  | ggccgagggtg | ggcagattgc  | tttgagctca  | ggagttcaag  | 166620 |
| aacagcctga  | gcaacacagc  | gaaaccccat  | ctctacaaaa  | aatacaacaa  | ttagctgggc  | 166680 |
| gtgggtggg   | acacctgtaa  | tcccagctac  | tccggaggct  | gaggggtggg  | aattgcttga  | 166740 |
| acctgggaag  | cggagggttg  | agttagccaa  | gatcgacca   | ctgcatttca  | gcctaggtga  | 166800 |
| cagagtgaga  | ccctgtctca  | aataaaaagg  | aaactactag  | cttgaagtat  | ccaggacacc  | 166860 |
| acctccccac  | ccccagctc   | cctgggtcac  | cccacctgga  | tgaccttgcc  | ctgggtgatcc | 166920 |
| agcttaggga  | cccccttatt  | cttaggggag  | tccagccctg  | gcataggaga  | aggcacacac  | 166980 |
| ttctctgggg  | acatgcacgt  | gcctctcct   | ctctgccaca  | gaatcagaca  | gtctggtttg  | 167040 |
| tgtcactatg  | gccacaaaga  | gcaagaggat  | aaaatatata  | cactggtcca  | tgtgctgtaa  | 167100 |
| aactgcctgg  | aggagcaaaa  | ttgctcccag  | ctggggccag  | actaatgggt  | gcctatgaag  | 167160 |
| atgggtgtgag | gtgggggtgg  | ggctaaggca  | gtccagaggg  | aggggtgggca | gaggctggaa  | 167220 |
| gcaggaaagc  | aggagcttga  | gccaacacct  | gcctggggct  | acctgagaga  | cagctccaag  | 167280 |
| gctcagcctg  | gagcctggga  | gggcaaggga  | gcccagcag   | gtgggcagggt | gggggtggcc  | 167340 |
| aggccagggt  | ctgggtctgga | ccacagtcag  | tggaggggag  | tgtcttcccc  | atgcagagga  | 167400 |



|             |             |             |             |             |              |        |
|-------------|-------------|-------------|-------------|-------------|--------------|--------|
| agcattggggg | ctgtggggga  | tgggggtgtc  | cctgctggct  | gcttgcta    | tctaagtctt   | 167460 |
| gcacttgcag  | aaacccgagg  | tccgagcccc  | ccagcaggcc  | tccttctctg  | tgggtgggtggc | 167520 |
| cattgacttc  | ggcaccacgt  | ctagtggcta  | tgctttcagc  | tttgccagtg  | accctgaggc   | 167580 |
| catccacatg  | atgaggtgag  | gtcggctggg  | ctgagagagt  | gaggtgggga  | gggtggggag   | 167640 |
| ttcctcatat  | cttgggtccca | aaagtactgt  | caccgagaca  | tggggctctcc | ttggaggccc   | 167700 |
| tgccaccccc  | agtctggggc  | tccccactgg  | ggtaaaagt   | ggaggatggg  | ccccaggcct   | 167760 |
| gggtccttgg  | cctcaactgg  | agcagctccc  | aaccctctgt  | aggaaaagtc  | agactttgtt   | 167820 |
| tttaatttttg | aatttttttg  | agatataaca  | aacataacaa  | taaaatatgc  | aaatgttttt   | 167880 |
| tatttttcctt | tttttttttt  | tgaggcggag  | tctcactctg  | tcgcctaggc  | tggagtgcga   | 167940 |
| tggcacgata  | ttagctcatt  | gcaacctcca  | cctctcgggt  | tcaagcgatt  | ctcctgcctc   | 168000 |
| agcctcccaa  | gtagctggga  | atacaggtgt  | gcgcactat   | gcctggctaa  | cttttgtggg   | 168060 |
| gttttttgtt  | tttttttgag  | acggagtctc  | actctgttgt  | ccaggctgga  | gtgcagtgc    | 168120 |
| gcaatctcag  | ctcactgcaa  | gctctgcctc  | ctgggttcat  | gccattctcc  | tgccctagcc   | 168180 |
| tcttgagtag  | ctgggactac  | aggcgaccgc  | caccacaccc  | agctaatttt  | tttttttttt   | 168240 |
| tgtattttta  | gtagagacgg  | ggtttcaccg  | tgtagccag   | gctggctcca  | aattcctgac   | 168300 |
| ttcgggtgat  | ccgccagcct  | cggcctccca  | gagtgcctag  | attacagatg  | tgagccacca   | 168360 |
| cacctggcct  | tttttttttt  | ttcttttgaga | cagggtctcg  | ctctgtcaca  | caagctggag   | 168420 |
| tgcagtgggtg | tgatcatagc  | tactgcagc   | ctcaaactcc  | tgggtccaag  | caatcctccc   | 168480 |
| acttcagcct  | cctgagtagc  | taggactaca  | gacagcacca  | ccacacotgg  | ctaatttaaa   | 168540 |
| aaaaaaaaaa  | attttttttt  | ggagagacag  | ggtctcactg  | tgttgccagg  | gctggctctt   | 168600 |
| aattcctggc  | ctcaagtgat  | tatcccactt  | tagcatccca  | aaaatgctgg  | aattgcaggc   | 168660 |
| gtgagccacc  | atgccagacc  | cacaaatatt  | tttctaagct  | tgtagacaaa  | cacacacata   | 168720 |
| tacaaacata  | tatgtgtttt  | tttaagtaaa  | agtgaatata  | acaacttget  | ttgttttaaa   | 168780 |
| gcttcaacac  | tattttgtag  | atatttcatg  | tcagcacata  | cgaagctacc  | actttctttg   | 168840 |
| aacggcctag  | tattccatag  | tacagatgtg  | tcataattta  | tccattcccc  | tattgacagt   | 168900 |
| tttagggggg  | ttttttcctc  | ctctcatitt  | tcactattac  | aaaaagtgca  | gcagtaatat   | 168960 |
| ccttcttttga | aaacttgtgc  | cgcaatctcc  | atggggtaga  | ttcctagagg  | tgattcctag   | 169020 |
| gtccaacatc  | tcaaactctt  | tttcttttcc  | tttttttagta | aaaaggaaaa  | gaaaaaataa   | 169080 |
| taaaacaaga  | aataaaaaa   | aaaacaagaa  | aatgaagggt  | ctaagggtcg  | agttgacaag   | 169140 |
| tcacacagtt  | gttttgaaag  | acatgtttca  | aaaatccatt  | caaagtgttg  | gttttatgtg   | 169200 |
| ggacotcaag  | gcacttgagg  | cctcaagacc  | ccagtcctct  | gaatatgtcc  | tagagggctcg  | 169260 |
| gcaggcagta  | ggtgaagcta  | caggagacct  | caaagctctc  | aaccagact   | gggatgggag   | 169320 |
| gtaagagggg  | cttcccaagg  | cctggagact  | ggatggaagg  | gtgaatggag  | acatagacag   | 169380 |
| acaggtaact  | gctcccaaaa  | tacatgggac  | ttgagcttca  | tgagactgga  | ctgtccttcc   | 169440 |
| ttccacatca  | caggaaacca  | caaaggctct  | gagcacttac  | agccctctag  | gagctgcagg   | 169500 |
| gcacccagaa  | tgtcttgcctg | atcaacacac  | atggctgagg  | ccatgctggg  | gctaaggaca   | 169560 |
| catccaggaa  | cactgtctct  | gtcctgtgga  | ggtgacagtc  | acaggaaata  | actgggggac   | 169620 |
| attccagcct  | cctcccaagc  | tctctcttcc  | caccagacac  | caagccccta  | acatctctag   | 169680 |
| aatacccatc  | tacttctctc  | catggccact  | gatcgtgccc  | aagttcagcc  | cgcaccttct   | 169740 |
| ctcacotgga  | tggctgcaat  | atcctcctcc  | ttggtttcct  | ttctccatgc  | tcccttccct   | 169800 |
| gcattctgtt  | gctccatggg  | agccagtgc   | ctttctttct  | aaaggcaaat  | gtggtcatac   | 169860 |
| ctcttggtta  | ctttaaatcc  | acgaagacag  | aattctgaat  | ggccttctgt  | gctcttcatg   | 169920 |
| atctcacctt  | agcattcttg  | tgttttctcc  | tgctgcctat  | ggagccgacc  | catagctagt   | 169980 |
| tctccctgct  | ctttcacatc  | tgtaagtttt  | tgcatatgct  | gttccctctg  | cctggattcc   | 170040 |
| cacccaagaa  | gggaggggtg  | cactgactgc  | cctgtgcctc  | cgcaggaaat  | gggagggcgg   | 170100 |
| agacccgggg  | gtggcccaacc | agaagacccc  | gacctgcctg  | ctgctgactc  | cggagggcgc   | 170160 |
| cttccacagc  | tttggtacaa  | ccgcccgcga  | ttactaccat  | gacctggacc  | ccgaagaggc   | 170220 |
| gcgggactgg  | ctctacttgc  | agaagttcaa  | gatgaagatc  | cacagcgcca  | cgggtgagtca  | 170280 |
| cagggtctca  | gacagggagg  | cggggccagc  | atggaaaagg  | gcagggtctaa | tgggggtggg   | 170340 |
| tgggacaaaa  | ccaaaacgtg  | tgaggaccgg  | cccgatggag  | tcgtggctga  | gagggggcgg   | 170400 |
| ggctaaaggg  | agacgtcgga  | ctccggtgtg  | ggcggagctc  | agaaatgagg  | tggaggcggg   | 170460 |
| gctaagtgtg  | gtggggctaa  | tagtgaagct  | ggggttgcag  | gaggggtggg  | gctaaggaga   | 170520 |
| ggggctgggg  | cagagctaat  | gtcacatggg  | gcaagagtgg  | gacggtggta  | aagaggaggg   | 170580 |
| gaagctccag  | gaaacggggg  | gatttttaaga | gcgaggtcgt  | caggaaatga  | gtgccaaagt   | 170640 |
| gaggcctcct  | gcagagccgc  | cctgtgtccc  | tgccaggatc  | tcaccttgaa  | gacccagcta   | 170700 |
| gaggcagtaa  | atggaaagac  | gatgcccgc   | ctggaggtgt  | tcgcccagtc  | cctgcgcttc   | 170760 |
| ttcagggagc  | acgcccctca  | ggtgcgctgc  | ggccccacct  | ctgcccagctg | tggcagggac   | 170820 |
| cccctatttt  | cccctcatcc  | gaaaccgctc  | ccccatcccg  | tccccgacat  | tggatgggta   | 170880 |

|             |             |            |             |             |             |        |
|-------------|-------------|------------|-------------|-------------|-------------|--------|
| gccaccgccc  | gagctcagag  | gtcatcttct | ccagtagcct  | cctccctttt  | tgtctggtag  | 170940 |
| agcctgcacc  | aagccatact  | gatgggaggg | gggcccattc  | ttccagctct  | gctggggaagt | 171000 |
| ccttcctgtg  | atlttgattag | tacctccagt | tcgcagaggg  | gctgaagacc  | accctccctc  | 171060 |
| caagccagct  | ttcctctcac  | tgcctccctc | tgtaccagga  | gctgagggag  | cagagcccat  | 171120 |
| cgctgccaga  | gaaggacact  | gtgcgctggg | tgttgacggg  | gcctgccatc  | tggaaacagc  | 171180 |
| cagccaagca  | gttcatgcgg  | gaggctgcct | acctgggtgag | gacgtgcagg  | cgggcccagag | 171240 |
| aacactgctc  | aggaagggcc  | aggcctgtcc | ccatgcttgc  | atgcacccca  | ccacccttga  | 171300 |
| gaccacagag  | tcattgtgga  | aagaacttca | gcctgctcct  | gatgggagtt  | tgtagagttg  | 171360 |
| ctcccaccag  | aagagggagt  | gggcctgcag | gaacagggga  | cagagggaca  | aaagacacag  | 171420 |
| cccagggcag  | tgtagacagt  | gccttaagtc | agggttccta  | aaagcagagc  | ccaagatgga  | 171480 |
| gagtcctttt  | tttttttttt  | tttttttttg | agacagtctt  | gctctgtcgc  | ccaggctgga  | 171540 |
| gcgcagtggc  | gtgatctcgg  | ctcactgcaa | gctctgcctc  | ccagggttcac | gccattctcc  | 171600 |
| tgcctcagcc  | tcccagtag   | ctgggactac | agggtcccgc  | caccacgccc  | ggctaatttt  | 171660 |
| ttgtattttt  | agtagagacg  | gggtttctac | gtgttagcca  | ggatgggttc  | aatctcctga  | 171720 |
| ccttgtgac   | cgcccgctc   | ggcctcccaa | agtgcctggga | ttacaggcgt  | gagccaccgc  | 171780 |
| gcccggccaa  | gatggagagt  | cttatatgtg | aggctctattg | aagaagggtt  | ctcaggagaa  | 171840 |
| aggcgagggg  | gggaggcggg  | acaggagaag | gagctatgaa  | ggatgtgggc  | tttgcctggag | 171900 |
| tctagcctca  | gggtttagcct | cagctagatc | ccatggggag  | ctgcagaggg  | agaactgtag  | 171960 |
| cacagagttg  | gggcccggctt | cttgacatc  | cgtatcgggc  | tgtcactggg  | tctggggaaa  | 172020 |
| tgggagggaa  | acctctccaa  | gtgaggccat | tccattcag   | ctgagggcag  | ttatccagag  | 172080 |
| gaggtagcag  | ctgtgagcca  | atagcagcca | acactcacag  | tggcaggcag  | ggcaccacaga | 172140 |
| acattcactt  | cacttggcat  | cagtattttg | agggacaccc  | ccaccaccgc  | ccacccttgt  | 172200 |
| tggttcttga  | gcaggggaaa  | taggactatg | aaataaaggg  | aggcgattct  | agccaatgca  | 172260 |
| tacaggtctt  | cctaccactg  | agatgctcaa | aggctctagag | caaggagttg  | gggggggctc  | 172320 |
| ccgtggagac  | agagagggca  | tagcctcaag | actcattgct  | tggctcgtaac | cctagagcct  | 172380 |
| tggatacaga  | ccctaagctg  | tccaaaaggg | atctcaagcc  | tgcctgtgga  | caccctggag  | 172440 |
| ggacccccaa  | tcctcccata  | cctatgggac | aggcatgacc  | cagggccagg  | tctgtaaatg  | 172500 |
| tctacacttt  | tcactggaag  | gagctaacc  | tgagcacagt  | ggacatccct  | gtgcaaagat  | 172560 |
| caggatggaa  | ccaaggctgt  | cccagccag  | aacaaacacc  | tgtccccc    | gcgtcccctt  | 172620 |
| gcctattttc  | ccaccctcct  | tctcacgctg | ttccctggct  | gctcaggaaa  | gtctggccct  | 172680 |
| gggtaaaagt  | ctccattcac  | acccttatcc | ctctccccta  | tctggtcaga  | atctggggaa  | 172740 |
| ccctacaaaa  | ccataatata  | cttcccacag | ggatccccta  | ccctgaaggga | aataactcca  | 172800 |
| gatctcaagt  | gtttttctcc  | accctaactg | cctgtcttgg  | tgcctggaat  | ttcagcttcc  | 172860 |
| tccctgctgt  | atcagattcc  | ctgggaacaa | agttctcctg  | gaagtggggg  | ggcctatggc  | 172920 |
| taccgtctga  | caccagccta  | gtggagaaa  | agggttcaat  | gagctaaaag  | ggctccccac  | 172980 |
| cacctcctct  | gagccatcac  | acaccccac  | attgtgctag  | agtctctgcc  | aataccccca  | 173040 |
| gccagcgtc   | agctggccgc  | ttgtacctgt | gaaggggaac  | cttgcctgcag | ccccgctatc  | 173100 |
| ttgggaaatt  | tgaaggggca  | gatcccaggg | gttctagggtc | tgcattctgt  | ctagagctct  | 173160 |
| ttcctgctgg  | gcaacttctt  | ggcatatttg | gccaggggcca | ttccctcccc  | agcctgggtca | 173220 |
| cagcgtggct  | atgaggcagc  | tccaaatttg | tgcagcacag  | aggggcctga  | gaggcctaac  | 173280 |
| atgggtgggg  | tgtgatggag  | gaggaggtgg | gagcccaca   | ggccgcgggt  | ccccacctca  | 173340 |
| cagtgccatc  | ttaggtgtga  | cagcccacag | tgcctgctga  | ccctgccac   | cacccatccc  | 173400 |
| caggtgggac  | tagtgtcccg  | agagaatgca | gagcagctac  | tcctgcgcct  | ggagcccagag | 173460 |
| gccgcctcgg  | tatactgccg  | caagctgcgc | ctgcaccagc  | tcctggacct  | gagtggccgg  | 173520 |
| gccccagggtg | gtggggcgcct | gggtgagcgc | cgctccatcg  | actccagctt  | ccgtcagggt  | 173580 |
| gagctgcccc  | cggggacacc  | accacccct  | ggagggtcag  | agggtcactg  | aagccagaag  | 173640 |
| ctcagccatg  | tctagtatga  | aggggagagg | gtacccacc   | tggaggagcc  | caatttgagc  | 173700 |
| aggcagaaa   | atggctgggtg | gagcctgtct | gaggaaggga  | ggcacagccc  | tggccaaagg  | 173760 |
| caacctcgtc  | tcagggtggg  | acatgcctta | cctagggcag  | cccaatctca  | gggtggagat  | 173820 |
| ttcaccttgc  | cctagggggag | caaagtctga | ggagggcagg  | actccgccc   | gccctgagtc  | 173880 |
| tgaggggttg  | gggagacaca  | gactacccca | ggaggaact   | gctgggggtac | tgacaaagga  | 173940 |
| ggaagtgcag  | ccactgcac   | agctgtcccc | ttcactcccc  | tcactctccc  | cgacacacat  | 174000 |
| cagccccaa   | ctcctgctgc  | caggaccagg | caccagggcc  | aaggacagct  | atgggtcatcc | 174060 |
| tcccaaatgc  | catctcccag  | ggcagggaga | agccctaagt  | cctgagtcct  | ctctgagact  | 174120 |
| ccaaagacct  | acctgcctcc  | gctcctccaa | acccctctaa  | ccctgatttt  | gccatgacct  | 174180 |
| gagacctcct  | gcgttaaagg  | aaggccctgt | gtccataaact | atttcccac   | agctgttgga  | 174240 |
| tacagggtgg  | gagtttgggg  | ttcaggattg | cctctccca   | gtcaggagca  | gggttgaggt  | 174300 |
| tcaggagcac  | tggctgctcc  | cagtgcctat | ggagggtcctg | ggcaggagga  | tgggagttga  | 174360 |

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| acgccatagc  | tggagcacct  | ccttctaate  | tcactccctg  | ctgtctcctg  | acccccagct  | 174420 |
| cgggagcagc  | tgcgaaggtc  | ccgccacagc  | cgcacgttcc  | tggtaggagtc | aggcgtagga  | 174480 |
| gagctgtggg  | cagagatgca  | agcaggtagg  | gggaaagggg  | gacggagtg   | tatccttggc  | 174540 |
| ccctaccggg  | caccatatac  | tgatgggggg  | aagggcagtg  | ttgcaaagcc  | cgtctcttcc  | 174600 |
| tcctccattc  | gctgtaccca  | acctggccgt  | ccccctacag  | tcacccgcac  | ccccaccca   | 174660 |
| ctcacagcgg  | cgcccctaac  | tcccactcct  | ccaggggatt  | ctccgcggac  | gctcgggtgg  | 174720 |
| agttgcagag  | cctctggaac  | catttctccc  | ccacaccctg  | cgcccatatg  | tggtaggtctg | 174780 |
| aggttcaagc  | acctgaagcc  | cctcacgtcc  | ctccccgcac  | cctgcagaca  | ggccttggga  | 174840 |
| cccggggcag  | ggctggaggc  | tgggcgaggc  | tggagggggc  | gcagggctga  | gggtgcgagg  | 174900 |
| ccgcccacga  | gtgtgtgccc  | gcgctcgccg  | ccgcaggaga  | ccgtacgtg   | gtggccgact  | 174960 |
| gcggcgagg   | caccgtggac  | ctgacggtgc  | accagctgga  | gcagcccat   | ggcaccctca  | 175020 |
| aggagctcta  | caaggcatct  | ggtgagtagc  | caggcggcgc  | cccggtagcc  | agcgcgaccc  | 175080 |
| gggctccggc  | cccgccactg  | ccccctggcg  | gcccggcgag  | cgctgacgcc  | ctcttcgccc  | 175140 |
| cctgctccac  | cccagggggc  | ccttatggcg  | cggtgggcgt  | ggacctggcc  | ttcgagcagc  | 175200 |
| tgtgtgtccg  | catcttcggc  | gaggacttca  | tcgccacctt  | caaaaggcaa  | cgccggcgag  | 175260 |
| cctgggtaga  | tctgaccatc  | gccttcgagg  | ctcgcaagcg  | cactgctggc  | ccacaccgtg  | 175320 |
| caggggcgct  | caacatctcg  | ctgcccttct  | ccttcattga  | cttctaccgc  | aagcagcggg  | 175380 |
| gccacaacgt  | ggagaccgct  | ctgcgcagga  | gcaggtgggt  | cctgagcccg  | cgggctcagg  | 175440 |
| caggggtttgc | cgaccgggga  | atgaccgtgc  | actggagggt  | cccgggcccc  | aaggaaacgt  | 175500 |
| gggggtctgc  | ctgattcatc  | ccacatatac  | actaagccag  | cagggcgctcg | gggtggggcg  | 175560 |
| gcgggggagc  | gcaggtgagt  | gccccagccc  | agcaggtccc  | acccacggaa  | tccgcagccc  | 175620 |
| gaactggggc  | aagacagaga  | atcatagcgg  | ggaggcgcca  | atgcctatct  | cctcccagcc  | 175680 |
| ttctctacac  | ccccaccccg  | ggccctgcgg  | gcccattgctc | ctcggtttcc  | ctgcacccaa  | 175740 |
| gcaaggggag  | gccccctcca  | ggacctcgta  | cctggaacct  | ggagcaggct  | ggcaactaaa  | 175800 |
| tcctctgagt  | gagtaggggtg | gagataaggg  | actaacatcc  | cgcaggtcca  | gtcctccaga  | 175860 |
| caccacgtgc  | agtcggtgcc  | caggcacttc  | tgcctggagg  | cagaggtaga  | gaataaggac  | 175920 |
| cacggacccc  | aaactggggc  | aagcagctgg  | gcccctgaccg | atggatattt  | gcccccttca  | 175980 |
| ccaccaacag  | cgtgaacttc  | gtgaagtgg   | cctcacaggg  | gatgctccga  | atgtcttgtg  | 176040 |
| aagccatgaa  | cgagctcttt  | cagcccaccg  | tcagcgggat  | catccagcac  | ataggtgagc  | 176100 |
| acctgagctt  | ggtccccac   | ccgcccctac  | atgaacaaac  | agatgcagaa  | taattccccc  | 176160 |
| ctatcagtg   | ctagatacct  | ccacacatcc  | atacactgtg  | atgagacct   | gaatcatcta  | 176220 |
| gaacacctgc  | gggatgaagt  | gcagtgggtga | ttaagagcta  | gagggttgat  | atgtagtctt  | 176280 |
| gccaaaggcaa | aaaacttctg  | gtgcctcagt  | ttccccattt  | ataaaatggg  | gtgatagtat  | 176340 |
| tgggtttctca | aaacgttatc  | acagggataa  | aatgagctga  | agtacctaga  | gtgagcacia  | 176400 |
| tgtcttgac   | acaatgtcta  | ggtgtttaat  | acgtgtaaaa  | tgcataatct  | tatctctcgt  | 176460 |
| cctccacgtc  | gtggtggggag | agaagtgggg  | agcgtgagtg  | ttggggaggc  | gaagccctcg  | 176520 |
| aggactcccg  | tgagctctca  | aagaaaagtg  | tcaaattggct | actttctagt  | cgccaggtag  | 176580 |
| gtacaggcta  | gggagggggag | gcgcgggtgg  | ccgcctagt   | gtggcctcag  | tggctctctc  | 176640 |
| tcctccgccc  | cttctcctct  | gcccccttca  | cccgcgtccc  | ccgctcctgt  | cccgagagg   | 176700 |
| ccctgctggc  | acggccggag  | gtgcagggtg  | tgaagctgct  | gttccctagt  | ggcggtctcg  | 176760 |
| ccgagtcagc  | ggtgctgcag  | cacgcgggtgc | aggcggcgct  | gggcgcccgc  | ggtctgcgtg  | 176820 |
| tcgtgggtccc | gcacgacgtg  | ggcctcacca  | tcctcaaagg  | cgcggtgctg  | ttcgccagag  | 176880 |
| cgccgggcgt  | ggtgcgggtc  | cgccgctcgc  | cgctcaccta  | tggcgtgggc  | gtgctcaacc  | 176940 |
| gcttttgtgc  | tgggcgccac  | ccgcccgaac  | agctgctggt  | tcgcgacggc  | cgccgctggt  | 177000 |
| gcaccgacgt  | cttcgagcgc  | ttcgtggccc  | ccgagcagtc  | ggtggccctg  | ggcgaggagg  | 177060 |
| tgcggcgag   | ctactgccc   | gcgcgtccc   | gccagcggcg  | cgtactcatc  | aacctgtact  | 177120 |
| gctgcgcggc  | agaggatgcg  | cgcttcatca  | ccgaccccgg  | cgtgcgcaaa  | tgtggcgccg  | 177180 |
| tcagcctcga  | gcttgagccc  | gcccactgcg  | gccaggacac  | cgcggcgcg   | cctccggccc  | 177240 |
| gccgcgagat  | ccgcgcggcc  | atgcagtttg  | gcgacaccga  | aattaaggtc  | accgcgctcg  | 177300 |
| acgtcagcac  | caatcgctcc  | gtgcgcgcgt  | ccatcgactt  | tctttccaac  | tgaggggcgcg | 177360 |
| ccggcgcggt  | gccagcgccg  | tctgcccggc  | cccgccctct  | ttcggttcag  | gggcctgcgg  | 177420 |
| agcgggttgg  | ggcgggggaa  | acgatagttc  | tgcagtctgc  | gcctttccac  | gccctccagc  | 177480 |
| cccgggggag  | ataaggtcat  | gggagagtgg  | gtggggacac  | acccagagac  | tggctttggg  | 177540 |
| attgggcact  | ggtccgctga  | ctgccaggct  | gaagggaccc  | gccaaggact  | gaacgggtaa  | 177600 |
| gagaagaggt  | ttgcaagaca  | gagcgcgcag  | cccggcaagg  | ggcatgtgac  | ccgaaggaa   | 177660 |
| gaacgcaaca  | gaagagtcc   | ggtctgaact  | tggccgagta  | ggggtggggg  | tgggatggca  | 177720 |
| ggaggagccg  | caggaggaag  | gaggttgtgc  | agggctgga   | cctgcagggc  | tgaagttcac  | 177780 |
| tcacgcaccc  | actcagcccc  | aaccgggagc  | caggcagaaa  | aacctgtgc   | cgtaggaaag  | 177840 |

|             |             |            |             |             |             |        |
|-------------|-------------|------------|-------------|-------------|-------------|--------|
| tgactggaag  | tggactccag  | agggacaggt | gtgggtggcac | agtcctgggtg | tggtgctgac  | 177900 |
| cacccaaata  | tgactgtgaa  | ttgtggaaag | ggcagtagat  | ctctaattgtg | gaggtgggaa  | 177960 |
| cattattgtg  | gtggaggcaa  | ttatgagggg | agcattttctt | tcgagacaaa  | acacccgtct  | 178020 |
| gggaaggccc  | caaggtcagc  | ttatgaagga | ccccacttgc  | acccaccccc  | agccatggaa  | 178080 |
| gagcagctgg  | aggggtggatg | gggaggccag | agggagcaat  | gaggggtggg  | cccagctctg  | 178140 |
| ctattgactc  | ggtatgcctt  | taggacattc | tcttaccgct  | catgggcctc  | agtttcctaa  | 178200 |
| agtgtgaaat  | gtcaggcact  | tccctctaac | tggcatgcaa  | cagccccacc  | tgcttgagag  | 178260 |
| ccctgaggtg  | acaataaaac  | atztatgtct | aaggggaagc  | cacagcctgc  | tgatatggcg  | 178320 |
| tggagaccct  | aatagtggga  | ggaatgcaag | ggttcccggg  | gctagagaga  | gaaggagaa   | 178380 |
| agctttcagc  | tgtgcatagg  | gaactgacca | gaaggggggtg | ctgctgtctc  | ccatcaagca  | 178440 |
| tcccaaacia  | ctccactgct  | taagacctct | ctggcctaca  | catgagggtcc | ctctctcctc  | 178500 |
| attcaaat    | attgtcttgg  | aagccagctt | ctggcctaaa  | atgccaccac  | ctgtgcatac  | 178560 |
| ctcttgtggg  | gctaggtgct  | ataataccac | gcgggtgcccc | tgctcctga   | gtgagtctac  | 178620 |
| ccaagtcttt  | ccctggccca  | tctgcaaagg | agtaggcatt  | accccaaccc  | cagagaacia  | 178680 |
| aaatccacct  | ggcctccggg  | atccactgga | agtttatctt  | tttaggggtt  | tatcccaacc  | 178740 |
| agtcgcttaa  | aaaccaagta  | acacagacct | gaggggtggg  | ggctgggggac | tgacacctcc  | 178800 |
| tcctactcat  | ggtggacagc  | agtggggact | agggaggggc  | aggagaggtg  | gctgaagcaa  | 178860 |
| ggcagcagta  | atggggccac  | gacgccacag | agccagctcc  | gtcctctccc  | agaccctggg  | 178920 |
| ggaggtccct  | gtggcttggg  | gtggggagtg | ggggaccac   | cccaggccct  | ccctctccct  | 178980 |
| tcctcagaca  | gctccttctt  | gggctcaacc | catttcttcc  | ggcaggagac  | tgaggcacac  | 179040 |
| agagaggagg  | aagtgggaga  | ggaggacgag | ggaggggcag  | ggtggcagca  | caaataagg   | 179100 |
| cagaggtgag  | aggcgtgggc  | aaggccactc | cacccccaca  | cccaccccag  | agagggggcg  | 179160 |
| ggaagccaca  | ccatcacgca  | gcatgtcggg | gggacaaggc  | ggggtttaag  | gctgaggggc  | 179220 |
| ccggggcagg  | cggggcctcg  | ggcctcagtc | aaagccgtgc  | cagtcgctgt  | gctctgagtc  | 179280 |
| gtattccagc  | tcggcgccca  | cacacttgac | accatccagc  | agcatgggcg  | tgccgtgggtg | 179340 |
| ccgggtctgca | aggcaggggtg | caagtacgtg | ccatgctggc  | ccccggcccc  | acccatgccc  | 179400 |
| cccactaagg  | ggacccctcc  | ccttccctca | gggctcagct  | ggaggtaggg  | acctgccaa   | 179460 |
| gaggttgaga  | acccctgagc  | cgggcaagga | tcccttgggt  | agccttgggt  | ccctgaggag  | 179520 |
| gacagaaacc  | ctcgcagtcg  | agcttgtgca | tccctctccc  | aaccaggagc  | ctcacgctct  | 179580 |
| cacccagac   | gcgggcctgc  | accgtcacgc | gcacacaggt  | gccagtgcc   | ccttcgcagg  | 179640 |
| ctagcagtat  | ctcctctttg  | aggacgccct | ctttgtgcgc  | cgagtactca  | cacttgacac  | 179700 |
| tataacctgc  | ccggggacag  | gggcaattgg | tcagcacctc  | ccccagcctc  | cccgatcctg  | 179760 |
| cccctggcac  | tcacgggccc  | ctaccaccca | ccatgggggg  | actgcatgct  | aagccccccc  | 179820 |
| aagaaagggg  | caggggatggc | ccttctggag | cccagagacc  | cattccccct  | agcaggggtg  | 179880 |
| cacaggtctc  | aaaaacccat  | tcttcagtga | gctggacatg  | cctccagcct  | gtggggccacc | 179940 |
| ccaatgtggc  | tccaagagtg  | aatgaagtag | ggtccaattc  | aggcttccaa  | aagaaggtct  | 180000 |
| ggctgttttc  | tcccaaaagg  | aaggcagggg | gagggcgtga  | tgaggagtga  | ggggggcagg  | 180060 |
| gcaggttagg  | ctttgagcag  | atccgatggc | aagaggtaaa  | ggcctgaggg  | gtgtcaatcc  | 180120 |
| attgaagggc  | aagggtcatga | agaggctggg | gacggggaca  | tgaaactgaa  | gctggagctc  | 180180 |
| cacgaaactg  | acacccatgc  | ctctgccagg | ctgggagtg   | ggcatggggc  | agggcctgaa  | 180240 |
| gtgagcctga  | taggaacagc  | agcctgaaac | ccaaggtctg  | ggactgggtg  | gtgctaggag  | 180300 |
| ttatccacac  | ggtgtgtgta  | gctccagcag | gtttttctag  | aggggttagg  | aggcaggagg  | 180360 |
| gaaggctgga  | ggcttcaaac  | cagttcctca | gcagctccca  | tcttggttac  | tgccccacgg  | 180420 |
| aggtaaccat  | cacaccatgg  | gcaggtagag | ggaagtatga  | gctcatggac  | ttctgttcag  | 180480 |
| gacagggagc  | aaggcctgga  | gtgtgggacc | tgccatgctg  | ccacagtgca  | agctcacaaa  | 180540 |
| gaggtgccac  | atccccgact  | tgctgagctg | cccatccacc  | ctagttagga  | gtaaggggagc | 180600 |
| accccatggt  | ctcactccca  | caccatcca  | ggccttctgt  | gaggagggga  | cgcaccttca  | 180660 |
| gggacgggca  | ccacgctgag  | gagcttgagg | tgccagctgg  | ggacaggtgc  | ctcgcggaca  | 180720 |
| tccttgctca  | gcctgtgcac  | tgggggcaga | gtgaaggtaa  | tctcatacct  | gtgcaggatc  | 180780 |
| ttcaggaagc  | caacctacag  | caggagaggt | gagggagggg  | ggcaccttca  | cttccctgctt | 180840 |
| cccacaggcc  | cacccagttg  | tgccacctat | gatcccatgc  | gccccagcac  | ctgcttcccc  | 180900 |
| agagagccca  | gaaagcccaa  | ctccctccat | cactccagct  | gatcgacccc  | cagccctttc  | 180960 |
| acagacccct  | tgaggggtccc | agccctacag | cttggccagc  | aagatcctca  | gcatctgtct  | 181020 |
| cacgggcccc  | aaagtccctca | gtggctcatt | tcatagacag  | gaaaatggaa  | accacaaaat  | 181080 |
| gactggccaa  | aggctatgta  | gtgtggccag | ggtcaagtcc  | tccctagact  | tcttggccat  | 181140 |
| tcctgctatg  | cttggcacc   | agcaaggcac | caccacatgc  | ctacaagggg  | aagttcaaa   | 181200 |
| tcccattggg  | catccacctg  | tcccatcac  | tgccagatgc  | ccagggtcag  | ggatgctggt  | 181260 |
| cctccattcg  | ctcataggat  | gacaccaggc | cacagatagc  | gtgggatgag  | ggatgggatc  | 181320 |

|             |            |             |             |             |             |        |
|-------------|------------|-------------|-------------|-------------|-------------|--------|
| aaattagaga  | tatgacagta | gaccagctgt  | gtcagcccag  | agctggccca  | cagacctggc  | 181380 |
| accagccact  | gaactggggc | tcctggctgc  | cctgctggca  | ctggactggg  | gatgcctgcc  | 181440 |
| tgccactca   | aaggctgtgt | ccaatggccc  | atactcacga  | cccaagcaat  | ttctgttggt  | 181500 |
| cattcacagt  | gcttggtgac | aaccacctc   | tgggacaaat  | gccattcttg  | gaaactccag  | 181560 |
| tagtatgaag  | ggtcacaagc | caggggtggt  | gctgagcagg  | ggggctggtg  | ggggtgccac  | 181620 |
| gccagcaggc  | aaagggtgcc | tgctgtactc  | ccagcttgca  | tgggcacaca  | gagtcctgct  | 181680 |
| cacagttact  | ggggctgggc | agcccatcc   | ctgggggcca  | actgggactg  | gctgcagaga  | 181740 |
| gttttagcca  | ttcaatggga | ccaggttgat  | attgctacat  | gacaaacttc  | aatccatgtg  | 181800 |
| ccccctcat   | gctctgctgg | tagggggcac  | atccctcaag  | aaggctttct  | ggctgtacct  | 181860 |
| gctaactgtg  | accctgtgac | ccaggggtac  | tactttataga | attttccagg  | ggaaatagag  | 181920 |
| atgtgcagaa  | aaatactctt | cgctgccta   | tttataacaa  | tgaaaacaaa  | aacagccgta  | 181980 |
| acaccagga   | acaggagacc | agctaagtca  | atgatgacac  | atatcatgga  | ccagtaggca  | 182040 |
| gttgtaaaat  | ctcgtggaaa | attatttact  | gacacaggta  | ggagacagtt  | cacaacagaa  | 182100 |
| gcaagcaaaa  | accagtgtgg | acacggtgag  | gccagctttt  | gtttttggaa  | acaaagcttt  | 182160 |
| ttaaaataag  | ctggatccat | caccatagtg  | ttttcagaaa  | aacaaataaa  | taaataaata  | 182220 |
| acatttttaa  | aagctggaat | aagccagggtg | tggtggctca  | cgctgtaat   | cccagcactt  | 182280 |
| tgggaggccg  | acacggacgg | atcacgaggt  | caggagatcg  | agaccatcct  | ggccaacaga  | 182340 |
| gtgaaacccc  | gtctctacta | aaaatataaa  | aaattagctg  | agagtgggtg  | cacgcgcctg  | 182400 |
| taatcccagc  | tactctggag | gctgaggcag  | gagaattgct  | tgaacctggg  | agctggagggt | 182460 |
| tgcagtgagc  | tgagatggcg | ccactgtact  | ccagcctggt  | gacagcaaga  | ctccgtctca  | 182520 |
| aaaagaaaaa  | aaaaagctga | aataacacac  | atcagaatat  | taacctgggt  | tattttggga  | 182580 |
| tagtatgact  | gaccgtacgc | tgaagaggat  | gcgtattgac  | tgtgctttcc  | ttctttgctg  | 182640 |
| gtggatatga  | aaactggaat | ggacattgct  | atgacgggat  | ggagctttct  | tacaaagctg  | 182700 |
| agcatgagtc  | ttaccataca | atctagcaat  | ttactccta   | gatgactacc  | caagtgaagt  | 182760 |
| gaaaacttat  | atccagggcc | gagcacagtg  | gctcacgcct  | gtaatcccag  | cactttggga  | 182820 |
| ggccgacacg  | gacggtcac  | gaggtcagga  | gatcgagacc  | atcctggcca  | acagagtga   | 182880 |
| accccgctct  | tactaaaaat | ataaaaaatt  | agctgagagt  | gggtggcacgc | gcctgtaatc  | 182940 |
| ccagctactc  | tggaggctga | ggcaggagaa  | ttgcttgaac  | ctgggagctg  | gaggttgcag  | 183000 |
| tgagctgaga  | tggcgccact | gtactccagc  | ctggtgacag  | caagactccg  | tctcaaaaag  | 183060 |
| aaaaaaaaaa  | gctgaaataa | cacacatcag  | aattattaacc | atggttattt  | tgggatagta  | 183120 |
| tgactgtggg  | tacttttaaa | tttcttcata  | ttttctctgt  | cttccaaatt  | ttctcatgtc  | 183180 |
| tattaagaca  | tgagtggatt | ttgcaatgag  | gggagagagc  | tatttttaag  | tgttggtttg  | 183240 |
| ttccatttac  | tctcttgggg | aagctgtcag  | ctgtaggaca  | aagagtggga  | acttcagagt  | 183300 |
| tggacagaaa  | tggatacaaa | cctgggcccc  | ccaggttctt  | ttttgtttgt  | ttgtttgaga  | 183360 |
| cagagtcttt  | ctggtgcccc | ggctggagtg  | cagtggcaca  | atctcagctc  | actgcaacct  | 183420 |
| cagcctcctg  | ggttcaagca | attcttctgc  | ctcagcctcc  | caagtagctg  | ggactacagg  | 183480 |
| cgtgcgccac  | cacgctcagc | taacttttgt  | attttttagta | gagacgggtt  | tcgacccatt  | 183540 |
| ggccagggtc  | aacctggggc | ccttttgagc  | agcgcagcag  | acccccact   | cagggccatc  | 183600 |
| tcaatggggc  | agacctccc  | cgacccaagg  | gcactcctgt  | tcaactctgg  | acccttgatt  | 183660 |
| tattgaccaa  | cctgaaggcc | tcagttccca  | ttttcccaaa  | tccagctcca  | ccaagcagaa  | 183720 |
| aacagggggg  | gataatacta | cttcaaagag  | cctaactgag  | ccagagattt  | ggcaaagaag  | 183780 |
| ggttaaaaaa  | aaagttgcac | cttgttattg  | ccatagttag  | cttcacacct  | gtatcacata  | 183840 |
| catacattct  | tccattctca | caccaactct  | aggggatgtg  | attgtctcca  | cttgagagaa  | 183900 |
| aagaaactca  | ggtgaccttc | ccaaggtcac  | agagccagaa  | tggctggcct  | agacctgaac  | 183960 |
| tcaggccact  | ggcaccacag | ccacgacact  | gccttccatt  | gctattgtct  | gggtgaagca  | 184020 |
| gatgacagca  | gtggctctgc | tcattctgat  | ctggatgaca  | ttctaaatgc  | tctcatctca  | 184080 |
| ttgaaatctc  | actgcctccc | atctgacaga  | tggggcaacc  | aaggcacagg  | gaaaggcaca  | 184140 |
| gacatgccta  | taagcccggc | tgagaggaac  | ggggataggc  | acccagaagc  | caggcagagc  | 184200 |
| cggggagggg  | aagaagctgc | tgatggggta  | ggtgtgcatg  | taccttgacc  | agaaagctgc  | 184260 |
| tgctactctc  | ctgggtgacc | atgaccaccg  | agtcatgcag  | cttctcatca  | aagtggacgt  | 184320 |
| ggctgtggga  | tccttctgca | tcgtggcctg  | ccgcaaagcg  | gatactccgg  | actctgggct  | 184380 |
| tgttgcctag  | gaagagtggg | agagggctct  | gagggctttc  | cctgtcccca  | ggaaactcct  | 184440 |
| cccccttgct  | ccctcgtcac | ccccaaagact | gcccttgaca  | tcattccagct | cccactggct  | 184500 |
| gggctctttt  | cagggctaga | tggacactag  | gatcatgagg  | cttgaggcct  | ccccgccggc  | 184560 |
| cccgacctgc  | ccctcccaca | aaaatccatc  | ccctgagagt  | gagcgtgggg  | ggactcaggg  | 184620 |
| actggctaata | atgacccttg | cttgaggggg  | gtggggctgg  | tgccagagcc  | aggaagggac  | 184680 |
| agtcttccag  | actcccacca | agccagggca  | gcgtgggact  | cagcccagtc  | ctctggatag  | 184740 |
| cctgcccagc  | gctctccttc | acagtgcata  | gctccccatc  | cctggggcct  | accctaagc   | 184800 |

|             |             |            |             |             |             |        |
|-------------|-------------|------------|-------------|-------------|-------------|--------|
| tgtgtcagtg  | catgaggccc  | ctgcctgccc | atcctcatga  | cccaaacctg  | aaagggcagg  | 184860 |
| gacaggaaca  | ggctctgggg  | gatctgcctg | gggtggcaga  | aacagggggg  | atcaaaaaaac | 184920 |
| acactcaggg  | cacctgggcg  | aggctgcagc | tgccactaaa  | cctcactggc  | ccctggcagc  | 184980 |
| aagaagagag  | aatccaagga  | agcttcacga | cccacctcaa  | gacccccctt  | cttcctctgt  | 185040 |
| ctagcccacc  | atcaaggccc  | tcagctgtct | ctgagctggg  | tcgggttctc  | caaggccaga  | 185100 |
| aagggcagac  | ggatgagaca  | ggcagggtag | gcgggagccg  | agctgaaggc  | ctggactgga  | 185160 |
| tgaatgctca  | ggctgtagag  | gccgagagca | cggccctagt  | cctctctgac  | ccctggcccc  | 185220 |
| taggccctcc  | caccaagagc  | cctgcccagg | gtcctcttgg  | cgggaggacc  | tgatccagct  | 185280 |
| ttgagccggc  | ctcttgata   | ctcggaggcc | caggggacca  | gcctggcaca  | tccacaatcc  | 185340 |
| ctgggtgagc  | ccacagaacc  | ctcaaatctc | accagatcac  | cctggcctaa  | aacctctcct  | 185400 |
| tatcccttct  | ggactctgaa  | agaaacccaa | actcccttcc  | acagcctgct  | gtggccctcc  | 185460 |
| cacatcacca  | gtctcagtg   | tcaccataag | cccctcggct  | cccccaaccc  | ccacagtgga  | 185520 |
| acggcagggg  | cacaagagca  | cacacctcca | cagaccagcc  | taaggagtca  | gcagggtctc  | 185580 |
| gcaggggtcc  | tgctagggct  | tgctgagcc  | ccagaacctg  | gcacaaacac  | tgacatagtc  | 185640 |
| ccaatcacat  | tcgcacacca  | gcacacacgc | actcacacat  | gcacacacac  | acacaggtgg  | 185700 |
| tgccggccct  | acctgcagca  | atggccctcc | taccagggcc  | caccggaaga  | gcttgtgcaa  | 185760 |
| gtcctaccac  | gccaggaagg  | cacttaccct | tggttggtgc  | agccatggac  | gccctccctg  | 185820 |
| ccacgcagct  | cctgccagac  | accgccactc | acgtcagca   | gcctcccatg  | ctccagggac  | 185880 |
| ccagccggga  | gcctgaagg   | tagggagtgg | ggagggcagg  | ggtcagggcc  | acctatggac  | 185940 |
| ccatggctca  | gggacaccag  | aggagctccc | tttaggaagg  | actcaaaccc  | ctaccgtctc  | 186000 |
| cagcagccca  | agtgggtgct  | gtactctcta | gcacgggggc  | ttctgcccac  | cctctaccga  | 186060 |
| ctgcccacat  | tcaccaagag  | ccctcacccc | ttcttgatac  | cccagtgaat  | cagggtcctg  | 186120 |
| ccagcatgtg  | ctgagctgca  | gcttctaggt | gccaaagcaa  | cagcatagga  | ctcctatccc  | 186180 |
| cagcaccccc  | agagactggc  | aggaggggca | gcctggaaag  | gaggacttta  | ttgggtatttt | 186240 |
| ccagggacca  | gttctgggtg  | tgctgacccc | aaaggctggg  | cctggagcta  | ccttatttcag | 186300 |
| gtcacacaag  | caatgcagct  | gggtgcggtc | agaggcaaca  | gggtgctcga  | tggtgggcaa  | 186360 |
| gaacagggga  | ccacataaag  | taagcgtgac | cttagagctt  | tccctcctgg  | tgatcggtca  | 186420 |
| gggccatctg  | caaaacagat  | gcctgtggag | ggggtgtgcc  | cccacctgaa  | ccccattcag  | 186480 |
| accctgcctc  | gagtcctcag  | cctgcctccg | cctcagctcc  | ccaacggcag  | cctcagcaca  | 186540 |
| ggggcagtg   | gaggcacccc  | agggaagcct | caatggggcg  | agctgggacc  | atctgagcat  | 186600 |
| caaaaagaat  | aatgagtgc   | actgatggaa | acagatggaa  | cacataaaaag | cagtgcattc  | 186660 |
| acagagatta  | ttagaaaaga  | agagagaaag | caaaacaaac  | aaacccaagc  | cctcaaaagc  | 186720 |
| cctcatttgt  | caccactgac  | ggtagcagtg | ccctctttac  | tctgaaagct  | gaggattaaa  | 186780 |
| gaggaagaat  | tcagcctgtc  | tcttgccctt | tcttgtaacc  | aaattgccct  | gggtggttgat | 186840 |
| ggaaagctct  | tctttcggga  | agaattcttg | ctaataaata  | acaaagaaat  | gaccaaattgc | 186900 |
| taagtcatte  | tgcaacccct  | aaaggaacaa | atgctggagg  | caacaagaac  | cagtggatgc  | 186960 |
| taaacagag   | gggaagggtg  | acaggaagca | gatactcaca  | ggcgcccaag  | cgtcactgac  | 187020 |
| agaggacctg  | gtatgggtttt | tggttttttt | tttttctgga  | gacggagtgt  | cactcttgct  | 187080 |
| accagggctg  | ggatgcaatg  | gcacgatctc | ggctaactgc  | aacatctgcc  | tcctgggttc  | 187140 |
| gagcgactct  | cccgctcag   | catcccgagt | agctgggact  | acaggcgccc  | gccaccacgc  | 187200 |
| ctagctaatt  | tttgatattt  | tagtagagac | agggtttcac  | catgttgggc  | aggctggtct  | 187260 |
| cgaactcctg  | acctcaagca  | atctgctcac | ttcggcctcc  | caaagtgttg  | ggattacagg  | 187320 |
| tgtgagccac  | catgcctggc  | cgaggacctg | gtgtttaaga  | ggaatggcac  | aagtgtgcaa  | 187380 |
| tgagggatat  | tgacttctcc  | cagtcacaca | ctgggtctatc | tctctactct  | ctgcaggaca  | 187440 |
| gcctgtcaca  | ctgatgcctc  | ctgaagtaag | gcagcccaaa  | gtcacagcat  | cactgacaag  | 187500 |
| ggattcctgc  | caaaaagggt  | taacctggaa | tctaaccag   | cctctaactt  | caacttctca  | 187560 |
| tttgctaaaa  | acacagagga  | gaggggaaca | aatttaatga  | ctccatgaga  | aagcaatgag  | 187620 |
| acgagccag   | aaggcaggat  | attctacagg | atgactgacc  | tggtttttgt  | ttttgttttt  | 187680 |
| gttttcaaat  | gtcaagaaga  | aaaagaaagc | aggctgggaa  | cgggtggctca | cgctgtaat   | 187740 |
| cccagcactt  | tcacggatca  | caaggtaagg | agttcaagac  | cagcctggcc  | aacatagtga  | 187800 |
| aacctcagct  | ctaccaaag   | tacaaagatt | agccgggtgt  | gggtggcgggc | agctgtagtc  | 187860 |
| ccagctagtt  | gggaggttga  | ggcaggagaa | tcacttgaat  | ccagaaggcg  | gaggttgacg  | 187920 |
| tgagccaaga  | tcgcgccacg  | gcacttcagc | ctgggtgaca  | cagcgagact  | ctgtctcaag  | 187980 |
| aaaaaaaaaa  | aagaaagaag  | aaaaataaaa | cgaggggagat | tattctagat  | aggaaagcag  | 188040 |
| caaaactacag | ctgattacac  | agctgatggg | aagggtgctg  | cctgttgtgt  | ttttgttttt  | 188100 |
| gtttttgaga  | caggggtctca | ctctgtccag | cccaggctgg  | agtcacagtc  | actgccaccg  | 188160 |
| taatctccca  | ggctcaagca  | atcctcccac | ccagcctgc   | caagtagctg  | agaccacagc  | 188220 |
| cacgcaccac  | catgcccggc  | taatttttgt | agagacagag  | ttgccagggc  | tggtctcaaa  | 188280 |

61/154

|             |             |             |            |             |             |        |
|-------------|-------------|-------------|------------|-------------|-------------|--------|
| ctccggagct  | taagegatcc  | acccacctca  | gcctcccaaa | gtgctgggat  | tacagacgtg  | 188340 |
| agccgccagg  | cctggcaaat  | aaatttttgc  | aaataaagtt | ttactaatag  | aacactacac  | 188400 |
| ttcattcact  | tacatggtgt  | ctgtttttgc  | actaaaacaa | agttacttcc  | aagtagttat  | 188460 |
| gatagagacc  | atatggcctg  | caaagtctaa  | accatttatt | atttgccct   | tcagagcaaa  | 188520 |
| agtttgccaa  | cccttggtct  | agatcaacac  | acacttaaca | gaccaaaaaa  | ggataatctt  | 188580 |
| caggacaact  | gacctgtttt  | tttcaccaca  | tcagctgcaa | gaagaaaaat  | aaaggagagg  | 188640 |
| gtgtttcgaa  | ctacctatca  | aagaagacat  | tttggggggt | gccgggcgct  | gcggctcaca  | 188700 |
| cctgtaatcc  | cagcactttg  | ggaagccaag  | gcgggtagat | cacctgaggt  | caggagtttg  | 188760 |
| ggaccagcct  | ggccaacatg  | gcgaaacccc  | atctctacta | aaaatacaaa  | aatcagctag  | 188820 |
| gcatggtggt  | gtgcacctgt  | aatcccagct  | actcaggagg | ctgaggcagg  | agaatcgctt  | 188880 |
| gaacctggga  | ggcggagggt  | gcagtgagcc  | aagatccgcg | cactgcactc  | cagcctgggt  | 188940 |
| gacagagcga  | gactctgtct  | caaaaaaaaaa | aaaaaaaaaa | aggccgggtg  | aggtggctca  | 189000 |
| cgctgtaat   | cacagcactg  | tgggaggctg  | aggcgggcgg | atcacctgag  | gtcggggagt  | 189060 |
| caagaccagc  | ctggccaaca  | tggagaaacc  | ctgtctctac | taaaaaatac  | aaaattagcc  | 189120 |
| aggcgtggtg  | gogcatgcct  | gtaatccagc  | tactcgggag | gctgaggcag  | gagaatcgct  | 189180 |
| tgaacccagg  | aggcggagggt | tgcggtgagc  | cgagattgcg | ccactgcact  | ctagcctggg  | 189240 |
| caaaaggagc  | aaaactccgt  | ctcaaaaaaa  | aaaaaaaaaa | agacatttgg  | gggatctgga  | 189300 |
| aaatgtaaat  | ggcctggata  | tgaggtaaca  | ctaagaaatg | agttgatcat  | ggaattgtat  | 189360 |
| tatgtaaaga  | attatgtcct  | tattttgtca  | gggaaacata | ctgaacaact  | tagagatgag  | 189420 |
| tgacactggc  | tcactcagtc  | aggacagaa   | gaggaactac | agggtggcct  | catagtttag  | 189480 |
| ggctgggct   | gtggagggca  | tgatctcaag  | gtccctctgg | tgccaagctg  | ctctgggagc  | 189540 |
| agatgtggcc  | tcacctcgct  | cccactttct  | caatgtgggc | aaagtccacc  | caggcccaag  | 189600 |
| ccctgcctca  | tgggcagggt  | gttcacagtt  | tgggtaccac | caggagcccc  | ctttggccaa  | 189660 |
| tggaccaggc  | cagcaacccc  | ctcttgagg   | tgagcttcca | ggccccagcc  | cagggtgcag  | 189720 |
| gaagtggagc  | tacacccatt  | cacctctggc  | caggggcact | gtggggtcag  | ttgcctcagc  | 189780 |
| cctggaaagc  | tcagcttgct  | cccagggaac  | ttggtcttgg | agagcagcct  | gccgtaaccg  | 189840 |
| ccccccgact  | acagctgccc  | ccaaagctgt  | gaactcagag | tatatgacaa  | atggccaagc  | 189900 |
| acaagagtgt  | gaggccctcc  | tatcctgcag  | aggctgctcc | agccacacct  | ctgcagccag  | 189960 |
| gaggcagtga  | cagcccagtg  | tcctcaaggt  | cccccatct  | cccaggatgc  | ttcctgaagg  | 190020 |
| ctcaggttaa  | actgcaccag  | tggctggtga  | gcaaaggctg | caggcccttt  | tcacaaaaca  | 190080 |
| ctctgaactg  | gtccatgcag  | ggcacctcc   | ctgtcttcc  | cccaacctac  | tgccctgaccg | 190140 |
| tccccccgcc  | ctgcctttgc  | ccaggtcctc  | ccttcagcca | ggaacaatcc  | cgctccctccc | 190200 |
| agccacacct  | tcgacagggt  | ctcagcccca  | acattttctt | tttaagggatg | cacacctgc   | 190260 |
| aaatcccagg  | acaggactac  | agctcttttt  | ggaatccct  | ttctcatgtc  | tgggatcatt  | 190320 |
| tgatatccct  | cttcctcacc  | ttacaagggt  | agtgaactgt | tttttgtcaa  | acctgtgtc   | 190380 |
| cccggcactc  | agcgcagcac  | cagacaggga  | ggagctgtgt | ttggtttatg  | tttgttgaat  | 190440 |
| gaatgaccac  | atcacatttg  | cttgaggggc  | ctggccaagg | cccgtacttc  | agccctaaat  | 190500 |
| gatattggtca | gggccagtg   | ctgcactggg  | ggctctgagg | aagcagaact  | ccccacccca  | 190560 |
| ctcttttgcc  | tcaccttgcc  | cctggggggc  | acttccctct | ccctggccccc | caccaagtgg  | 190620 |
| cctcccactg  | gcactcctca  | ggccttgctg  | cagtcagctg | gttacctgtc  | catgcctgct  | 190680 |
| ttgagaggga  | catgccccat  | gcccactgag  | ggtaggacc  | acatctcatg  | cagtgccagt  | 190740 |
| actcagtaaa  | gggcaaacac  | tgagggcctg  | taacctctg  | gatagtga    | acatagaggc  | 190800 |
| aggaagcaag  | ggacttcagg  | aacccaaagg  | aaactgggaa | aaccaaacct  | cctttctcaa  | 190860 |
| tggagaacct  | gctggctgct  | gtcctcgggg  | agctagactc | ttgtgcacac  | acaattcctc  | 190920 |
| atcagaaaca  | ggcaaagagg  | gcaactgggc  | tccctgctg  | cagctgtgtc  | cacagagaac  | 190980 |
| agctaggcca  | ccacacaacc  | ccaagcctgg  | ccctgcctcc | agcctcctcc  | tgtgccaacc  | 191040 |
| aaggcctcac  | catggctcaa  | tccacatggc  | ccccagaggc | agcagtcctg  | gttcaaattc  | 191100 |
| aggatctgcc  | cttgctagct  | gcatgggtga  | gttctttcac | ttctctgtcc  | ctgtttccta  | 191160 |
| atgtgtaaca  | caggaataag  | cagtggcctc  | ttccttcag  | ggtttgctgc  | aactgtgcct  | 191220 |
| gaagctctgt  | gacaatgcac  | cccccatctc  | tgcaaaagca | aacccccaaa  | ggcctgggtt  | 191280 |
| ctaaagcaac  | agcagtttca  | gcagcacctg  | cagagaggca | cttcaggcca  | atcctggagg  | 191340 |
| agccaggagt  | gacccttaga  | gtgggcccc   | gggacgtcag | ccttcttgga  | aaacagctca  | 191400 |
| aggggtgagg  | gggcctccct  | cctcctgcct  | cccccttctc | ccactcccaa  | agcagccagg  | 191460 |
| tccctaggga  | gggtcagaga  | acagatgctg  | ggagtttcca | gtccccctaa  | ccagaggggg  | 191520 |
| tcacaaggaa  | gatgtgcaga  | atgaacatcc  | tgggaaactg | ggaaatgact  | agggaggaac  | 191580 |
| atggtgcctc  | ccccccagca  | aaaaaaatta  | tacccttccc | catgagatgg  | agtgtcagca  | 191640 |
| agcttccagg  | ccccagccca  | gggtgcagga  | agaggagcta | caccatttca  | cctctggtca  | 191700 |
| gcatgctccc  | aacgtctgag  | acctcatttc  | tcattccttc | ttcagtcctc  | attctcattg  | 191760 |



|             |             |            |            |             |             |        |
|-------------|-------------|------------|------------|-------------|-------------|--------|
| tggttcgagg  | tctttctctc  | tgagccagga | actccgcaac | cttccaccct  | ccctacctct  | 191820 |
| tcctctccta  | ccccagcctg  | gggctcagtc | ctggctcaag | cactcagtc   | agcagaagca  | 191880 |
| ctgtgtagcc  | tcccattaaa  | gctcacgcct | gtgaaaagaa | cacccattga  | ggccttgaga  | 191940 |
| tggtggccaca | ctgacccgct  | gactctcagg | actggacaca | gcagaggcca  | cacatactca  | 192000 |
| gaacaaagcc  | tggaaaggca  | aggctggagg | tcagtagttg | tggcagcttc  | acatcaactc  | 192060 |
| agctttaatg  | tgatttaatt  | tccttctccc | tccagtgggc | caaaggtgca  | aagataagta  | 192120 |
| tggtctgttct | ctctccttct  | aacagtggag | tgctgggggt | gggggtgggg  | gaatatggag  | 192180 |
| aagggaacct  | caccacccac  | accttctctg | ctccccaaca | agtgtctgcc  | tcctctgccc  | 192240 |
| agcattctcc  | ccactttgcc  | ctcagctagt | gggtgcttag | cctccagata  | gcctgcccc   | 192300 |
| cctaggccct  | gccctgggcc  | tgtgatccag | aggtcccaag | aagcagaggc  | caggctggat  | 192360 |
| ccaggggggtc | agccaagggtg | agggtgggag | cacacaggat | tatctcccag  | ggacagggt   | 192420 |
| gctgcctcgt  | agctcaggat  | ggatagaatg | tggggggata | tccagctaca  | ttttccctcc  | 192480 |
| acaaaagacc  | agaatgggag  | ggggatgggg | tgctgccccg | actttcttca  | actccccgga  | 192540 |
| gcagaaaaat  | gccctacctc  | cactttccag | tgccaagatt | caagaagaaa  | ggcaagcgga  | 192600 |
| gacttccctt  | tctcagtcct  | tgcttactaa | tggaaacacg | ggtccagaac  | ctaaatccag  | 192660 |
| tcctctctcc  | ttcataccac  | cgggagggag | gtgcagccca | agcccccgag  | gccccaaagg  | 192720 |
| tccaggtgta  | ggacctttta  | tcctctccgg | cagccatccc | tgtgggtgtg  | gcacccccgc  | 192780 |
| cacaccccat  | tcttgtcatc  | ttagggggag | gggggaaatg | taatcggaca  | tcacccccat  | 192840 |
| ccaattccatc | ctgagctcgc  | aggcggcgcc | tctgtccctc | ggagataatc  | ctgtgactc   | 192900 |
| cccaccttca  | ctcacctcgg  | ctgacgcagg | agtctccgga | gcccgcactc  | ccagcaatca  | 192960 |
| ctgccctcct  | tcctgagggg  | gctgaggctc | ggaggctcag | agatgctact  | ggtccaaggt  | 193020 |
| catgcagcga  | ggcagcggca  | atgaacaggg | tcgtggggag | gagggggcgc  | tgacccccatt | 193080 |
| acgccccgc   | cctcactacc  | gcactcatgc | ccccggacaa | tcgttctcgc  | gaaaacaccc  | 193140 |
| cagctgccac  | cagttaacgg  | tcactgcgcc | ccgggaacct | gacatcactc  | tagttcaggt  | 193200 |
| gctggggagt  | cttccaggcc  | cggcccccaa | cagcggagcc | ccccaccccc  | agccctgcca  | 193260 |
| tcacggccgc  | tgggggtcca  | ggcactgact | gctcaggaca | gggcccgggac | cggggagggg  | 193320 |
| gacctcggcg  | aacgggaagg  | aaccgggagg | cagggagcag | aggggtggcc  | tcaccttgcc  | 193380 |
| gggtgcaccc  | cggggccggg  | gagggcaggg | acgaccactg | cagcggcgcc  | ggctgcagga  | 193440 |
| gctcaacgcc  | gagcacgagg  | aagggagccc | cgcgcgcgg  | ccgcccctcc  | gtcggcagcg  | 193500 |
| ccccgcctcc  | gcccattggg  | tgatctggga | gggtggggcg | agggacgctc  | cggaccaatg  | 193560 |
| agcgggctcc  | aaagaacggc  | caactggcga | gggcccgcct | cgtcacgtgc  | cagggtcgcc  | 193620 |
| gaggcagcgc  | cctgctagtc  | cgcgcctgcc | gggcgagctc | tcgcgaggaa  | gacgggcagg  | 193680 |
| cggcccaact  | aggccagggg  | ccagaaccga | ccactcgaag | agggagaagg  | agggcctcgg  | 193740 |
| ataggccccc  | cccccgctcc  | ttcttccgcc | tgggggatag | cgcctctagc  | cttgaacctt  | 193800 |
| gcttaggacg  | cacctccctt  | gggcccctcg | ctctcgggag | ggctgtcggg  | cgcgtctcgg  | 193860 |
| ggctgggtgg  | agctcccgaa  | ggtggccttt | ctccctgggc | ttccacgcgc  | gcttcggcca  | 193920 |
| tcgatacggg  | cgtgttggg   | ctcgttcagg | agctgaggaa | ccctccatca  | ctcctgtttc  | 193980 |
| gaccccagg   | tttggacctc  | ttccccttcc | accccatccc | ctgtcttgaa  | agaagcaacc  | 194040 |
| cccgctcggg  | cccgagacgc  | gtcccgggtg | cctggccggg | cctggagaag  | catcagaaca  | 194100 |
| aagaaggcac  | gcgggctggg  | ggctgggaga | gcctgtgacg | cgcccccggg  | gaccgcagcc  | 194160 |
| tctgtctccg  | gtctccatgg  | aggcggctcg | catggcagcg | agatgcgcct  | cgtcagcac   | 194220 |
| cgcgggggtg  | gatgtgggcg  | cctgcaatga | gccgaggagc | gagaggcgtg  | gccctccggg  | 194280 |
| ctgcgggggt  | tcttgccggg  | gctctccgcc | cgcgcggttc | cgaacacccc  | acctatactt  | 194340 |
| cgcgcgtggg  | gacggattcc  | ccaaagtgcc | ctcagtaagc | cgtccggagc  | acgcagcgcc  | 194400 |
| ttgcttccaa  | cggaaactaga | gagacggcct | gggcggccga | aggccagcct  | cccttcagca  | 194460 |
| gggcgggggt  | cgtgccttta  | aaggagccct | caagtctgcc | accctgtggc  | ccataacctg  | 194520 |
| tctgctgac   | tccagtctgc  | acactgttgg | caaattaatc | tttctgagct  | cttgttttca  | 194580 |
| tcgcgtccct  | ctcctgctcc  | aaagccctct | gggactgcct | ccagtagcgc  | ttcacaaact  | 194640 |
| tcagcagcac  | tttgggtgac  | tcatgtgccc | tcgcgtttga | gagacagcgc  | ttacattgag  | 194700 |
| agctttttcac | attctgatct  | cagcccatcc | actcctgccc | actctaccca  | ctcccaccac  | 194760 |
| actaccttta  | gaccttgtct  | tgaaagtcta | aagttggcag | acgggtgggc  | caaagtggcc  | 194820 |
| cgtaggtgta  | ttttgtttag  | ctcacatggg | gctttttaaa | agaagtcctt  | ccccctcctg  | 194880 |
| cctcctccgc  | ctctttctgg  | aatcaggaat | ctggtttcgt | ggcttttgaa  | atatcagaag  | 194940 |
| atctgacaac  | gtgggctcac  | tttccagcct | ggcattacaa | gcccccttaa  | gagcacataa  | 195000 |
| atttaccata  | gtccccgcta  | atccctgttt | ctcacggtta | gttaccttct  | gggtctgtgc  | 195060 |
| agacatttgt  | gatccccctc  | cctcatttat | cccaccttcc | aacctggaag  | gccctccccg  | 195120 |
| ctctcttctc  | atagccaaaa  | ttcaagccct | cttcaaccaa | gataagtttt  | gcctatttat  | 195180 |
| ttattttatt  | atttttgaga  | cagagtctca | ctctgtcgcc | caggcttgag  | ttcagtggtg  | 195240 |



|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| ccatctcggc  | tgactacagc  | ctccaccccc  | tcccccgccg  | ggttcaagcc  | attctcctgc  | 195300 |
| ctcagcctcc  | caagtagctg  | ggattctagg  | tgcgcgccac  | cacaccccagg | gctaattttt  | 195360 |
| gtatttttag  | tggagacagg  | gtttcaccat  | gttggtcagg  | tgggtctcaa  | actcctagcc  | 195420 |
| tcaagtgatc  | tgccctgcctc | agcctcccaa  | agtgcctggga | ttacaggcat  | gagccaccgc  | 195480 |
| gcctggccca  | agcctgcctt  | tttcaaaagt  | ctttctagcc  | aacttctcaa  | aaccatctct  | 195540 |
| gggtgtgggt  | ctctcaaaca  | caaactggct  | gctaactgca  | cagcccgggg  | ttatttccga  | 195600 |
| attgtcaact  | cctttcagca  | aacatttact  | gaatgtatct  | gatgagtaag  | ttattgtgct  | 195660 |
| aggctctgtg  | gagtgggaag  | cactcacagt  | ttagtgtagg  | agggacacca  | cacacaaaga  | 195720 |
| actgtggtct  | cctgtccttg  | ccaaatgttc  | ctcacactcc  | caggcatcgg  | gcagtctggc  | 195780 |
| ccactgggag  | gagctgaaat  | acagagctgc  | ccatgcctag  | ggatactaca  | tgctctcagg  | 195840 |
| gaaggccaaa  | gatggttctg  | gaaagcctcc  | aggagaaaat  | tagactcttt  | gtgtcaggag  | 195900 |
| gaagtcaagg  | caggagttac  | ctggctgccc  | catcactgcc  | caatgtgtct  | gtgttaccac  | 195960 |
| acgagtagtg  | cccaggctct  | ggggcttctt  | gtgcggaggg  | ccttaccaaa  | gatagatggc  | 196020 |
| tctgcaaggg  | aaaccacagg  | cccctaccag  | gccccctgga  | gacagaggca  | ggacaagaag  | 196080 |
| aatcaggaac  | aactccaggt  | taggagaaca  | gatggagcac  | gtataaaaca  | tccactcacg  | 196140 |
| gccgggctg   | gtggctcacg  | cctgtaatcc  | cagcactttg  | ggaggctgag  | atgggcggat  | 196200 |
| catgagggtca | ggagatcaag  | accatcctgg  | ctaacacggg  | gaaaccccgt  | ctctactaaa  | 196260 |
| aatacaaaaa  | attagccggg  | catgggtggc  | gatgcctgta  | atcccaacta  | ctccggaagc  | 196320 |
| tgaggcagga  | gaatggcgtg  | aaccccggga  | ggcggagcct  | gcagtgagcc  | cgagattgcg  | 196380 |
| ccactgcact  | tcagctttgg  | gcgactgagc  | cagactccat  | ctccaaaaaa  | aaaaaaaaaa  | 196440 |
| aaatccactc  | accatgagaa  | cctgggtgact | tttaaaaaaa  | aacaaaaaaa  | cacaacactt  | 196500 |
| gctacaactc  | caaacacacg  | gttttttgtt  | ttttttttta  | tttttcagag  | acagggtctc  | 196560 |
| actatgtttg  | ccaggctggg  | ctcaaaactc  | tggcctcaag  | cgatcccccg  | accttgggct  | 196620 |
| cctaattgtg  | tgggaccaga  | ccagaggcac  | aagccactgt  | gcccagccct  | ctattctttt  | 196680 |
| taattaaaaa  | ttaaacaact  | ataacctcat  | tggaaaactg  | ccttggcagta | cctaataaaa  | 196740 |
| atgataaatg  | aaatgagtaa  | ttccactttt  | tttttttttt  | tgagacagag  | tcttactctg  | 196800 |
| tcacccaggc  | tgggaatgcag | tgggtgcgatc | tcagctctct  | acaaacctct  | cctcctgggc  | 196860 |
| tcaagcactc  | ctcccacctc  | agcctcccca  | gtagccagga  | ctacaggcac  | atgccaccac  | 196920 |
| accagctaa   | tttttgtatt  | tttttttttt  | tttttgagatt | acaggcatgt  | gccaccacgc  | 196980 |
| cggctaattt  | tgtattttta  | gtagagacgg  | gtttctccat  | gttggtcacg  | ctagtctcga  | 197040 |
| actcccaacc  | tcaggtaaac  | tgccctgcctc | agcctcccaa  | agtgcctggga | ttacaggcat  | 197100 |
| gagccaccgt  | gcctggccaa  | tttttttgtg  | tttttaatat  | agacgggggt  | tcaccatgtt  | 197160 |
| ggccaggctg  | gtctcgaact  | cctgacctca  | ggtgaccgac  | ccacctcaga  | ctcccaaagt  | 197220 |
| gctgggatta  | caagcttgag  | ccaccatgcc  | cagccatggc  | atgtgtttta  | ttttaaaaga  | 197280 |
| aactggcaaa  | ttgtttttcca | aatgactgt   | acttacttac  | actcccatta  | gcagagtata  | 197340 |
| agagttttct  | ttgcttcaca  | tcctcatcaa  | cataaaaatt  | tgcaattttg  | taaaaaacgc  | 197400 |
| tacaccatgc  | taaccaaaca  | aagcagggca  | acggactacc  | agtcttcagc  | ctctcctctg  | 197460 |
| agccatgccc  | ttgtagagga  | caatttcaca  | acagctatca  | agattacaaa  | tgcacacaca  | 197520 |
| aaaaaagaga  | aaacacacac  | acacacacac  | acacacacac  | acacagccaa  | catgggtgaaa | 197580 |
| ccccatctct  | actaaaaata  | caaaaaatta  | gctgggctg   | gtggcagggt  | cctgtaatgc  | 197640 |
| cagctaccca  | ggagggttgag | gcaggagaat  | cgcttgaacc  | cgggaggcag  | aggatgcagt  | 197700 |
| gaggcaagat  | agtgtcattg  | cactccagcc  | tgggccacag  | cgcgagactc  | tgactcaaaa  | 197760 |
| aagaaaaaag  | agagaggaaa  | aaaaacacaa  | caaagattac  | agatgcatat  | attctttgat  | 197820 |
| ctaggaaccc  | cagttctgga  | atttatcatt  | tagatatact  | tccccatgat  | actaaatcat  | 197880 |
| gagtgtatga  | ggttattttat | tgttgcattc  | ttgtgatagc  | aaaagatttg  | aaataactca  | 197940 |
| aatgtccata  | aatagaagac  | ttcttaataa  | aatttggtag  | agacaatata  | atgttatatg  | 198000 |
| gcttttaaaa  | gaaaaaaaaa  | aaagagggaag | ttttcaatgt  | cctcccaagt  | agccaggatt  | 198060 |
| acaggcgtga  | gccaccatgc  | cgggataatt  | tttgtatttt  | tagtagagat  | ggggttttgc  | 198120 |
| catgttatac  | aagctgggtc  | cgaactcctg  | gcctcaagcg  | atccaccac   | ctcagcctcc  | 198180 |
| caaaatgcta  | ggattacagg  | catgagctac  | cgtgccacgc  | ctgctaagga  | aagattttgt  | 198240 |
| attagctagg  | attctccaga  | gaaacagaac  | taggaaagag  | ggcagatagg  | aaagactgac  | 198300 |
| atattataag  | gaattgactc  | acacaattat  | ggagtctgac  | cagtcccaag  | agctgcagg   | 198360 |
| tgggttggca  | agttggagac  | ccaggagagc  | caatggttta  | gttccagctc  | gagtcccaat  | 198420 |
| gcctgagagc  | caggaaaagtc | aatgatgtat  | ctccagtcga  | aagggcagca  | ggttttgagc  | 198480 |
| caaggaagaa  | tcaatgtttc  | agtttgatcc  | caaaggcgag  | aaaaaagctg  | atgttccagt  | 198540 |
| tcaaagacca  | cagggaacaa  | aaaattctct  | cttatttggg  | gagggccagc  | ctttttcttc  | 198600 |
| tatttaggtc  | tcaactgact  | ggatgtggcc  | cactcacatt  | agagagggaa  | atctgtttca  | 198660 |
| ctcatctgct  | gattttcttc  | tttttctctt  | ttgtttttga  | aatggattct  | cactctgttg  | 198720 |

64/154

|             |             |             |             |             |             |        |
|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| cccaggctgg  | agtgccgtgg  | catgatcttg  | gttccctgca  | acctccacct  | cccgggttca  | 198780 |
| agcgattctc  | ctgcctcagc  | ttcctgagta  | gctgggggta  | caggcaatca  | ccaccacgcc  | 198840 |
| cggctaattt  | ttgtattttt  | agtagagacg  | gggtttcgcc  | atggtgacca  | ggctggtctt  | 198900 |
| gaatgcctga  | cctcaggtga  | ttagcccgcc  | tcggcctccc  | aaagtgtctg  | gattacaggc  | 198960 |
| gtgagccacc  | gtgcccggcc  | tcctctgctg  | atttaaattgt | taaccttagc  | cgggtgcggt  | 199020 |
| ggttttacccc | tgtttttccc  | gcactttggg  | aggccgaggc  | gggcggctca  | ctcgagctca  | 199080 |
| gggtgttcggg | tcaacatggt  | gaaacctgcc  | tctactaaat  | acataaaaaat | tagctgggca  | 199140 |
| tgggtggcatg | agcctgtagt  | cccagctact  | caggaggccg  | aggcacgaga  | atcacttgaa  | 199200 |
| cccaggagggt | ggagggtgca  | gttaactgaa  | attgtgccac  | tacactccag  | ccagggtgac  | 199260 |
| aaaacaagac  | tcttctctca  | aaaataagag  | taagacaaat  | gtgcatgttt  | gtttatagtc  | 199320 |
| acataaagtc  | attccgggaag | gatacataaa  | aactaaagat  | agagggtccc  | tattagagat  | 199380 |
| gtgctgggaa  | taagaagact  | tttcaactttt | tatgtgtttt  | ttttaaatca  | tatatatggt  | 199440 |
| ttatctatct  | aaaacattaa  | gtaacaaagt  | ttaaataccc  | ttgcatgcct  | tcctctcctg  | 199500 |
| taggggttgtg | aacactggag  | ggaacaaatc  | atgtttttata | ctcacttttg  | tattcctggt  | 199560 |
| atlttagtaca | tagtaggtac  | tcaataaata  | gaggaaagaa  | cgaatgaatg  | aaccagaaga  | 199620 |
| tagcagattt  | agtgtcgggt  | agagagagga  | gaggggccag  | gcacagtggc  | tcacacatgt  | 199680 |
| aatcccagca  | ttttgggagg  | ctgaggcaag  | aggctcactt  | gaacctagga  | gttgggagacc | 199740 |
| agcctgggta  | acaaagcgaa  | atcctgtctc  | tccaaaaaaa  | aaaaaaaaaa  | aaaaaaaaaa  | 199800 |
| aaaagcccg   | gagcagtggc  | tcacacctgt  | aatcccagca  | ctttgggagg  | ccgagggtggg | 199860 |
| gtgatccacc  | tgaggttagc  | agttcgagac  | caactaggcc  | aacatgggtga | aacaccatct  | 199920 |
| ctactgaaga  | tacaaaaaatt | agttggcgct  | agtggcagggt | gcctgtaatt  | ccagcaactc  | 199980 |
| aggaggctga  | ggcaggagaa  | tagcttgaa   | ctgggaggca  | gaggttgag   | tgagccgaga  | 200040 |
| tcctgcccgt  | gcactccagc  | ctggatgaca  | gagtggagaca | tctcataaaa  | aaaaaaaaatt | 200100 |
| agctggatat  | gggtggtatgc | acctgtcctc  | ccacctattc  | tagaggctga  | gggtgtaaaga | 200160 |
| ttgcttgaac  | ttgggaggct  | gggttgagca  | tgagccaaga  | ttgtgccact  | gcactccagc  | 200220 |
| ctgggagaca  | gagtaaaacc  | ctgtctctaa  | aaaagaaggg  | agttagggct  | gggtgcgggtg | 200280 |
| gctcacggct  | gtaatcccag  | cacttttagga | ggccgaggct  | gggtggatcac | ctgagggtcag | 200340 |
| gagtttgaga  | ccagcctgac  | caatatgggtg | gaaccccgct  | tctactaaaa  | atacaaaaag  | 200400 |
| tagccggggtg | tggtggcggtg | tgctgtta    | ccagctact   | tgaggaggctg | agacagagga  | 200460 |
| atcacttgaa  | ccagggaagt  | ggagagtga   | gtgagccgag  | atcataccac  | tgcactccag  | 200520 |
| cctgggtgac  | agagcaatac  | tctgtctcaa  | aacaaaaaaa  | aagaaagaaa  | gaaaaatttt  | 200580 |
| ttaaaaatta  | aattaaaaat  | aagggtggtta | taatcttata  | taatgtaata  | acaaaagtgtg | 200640 |
| acatcctatc  | acatttgctg  | tattctactt  | atcagaagca  | agtcactaag  | tccagcctac  | 200700 |
| attcaaggaa  | aggggtacaca | tggtgggtgac | gtccaggagg  | cagggattac  | tggtgagccac | 200760 |
| gttaggagct  | gtctactaca  | aggaggaaaa  | agcaatgaaa  | ggatgaaatc  | taggatgggtg | 200820 |
| gttatctctg  | agtggaaagca | agggggagag  | atgcaagaag  | cacatggccg  | gtgaaagtta  | 200880 |
| tttgtaataa  | ttccattctt  | ggggcagggtg | gtggatttat  | aggtgttcat  | aagtgtgttg  | 200940 |
| attatggaca  | aatggggaga  | gagtgtcatg  | aaccaaggat  | tatatattagt | ccaattcagt  | 201000 |
| aatattttaa  | atattcaata  | ataaatgaaa  | taaaaattgc  | caactactcc  | tgacctgtgt  | 201060 |
| ctgagcccca  | tccatgcact  | ccacctgggtc | ttagttaa    | actggctcca  | cctgacctcc  | 201120 |
| cacttgcccc  | aaactaaccg  | ttagcccttc  | acacacactg  | acccttgcta  | tagcaccacg  | 201180 |
| atctctccat  | agtctgcaat  | aggcccccg   | agttctattc  | cctgaaactt  | ttatccatgt  | 201240 |
| aacctcacct  | gtaggctgct  | atatgtaaag  | atagtaggaa  | tggtgatgg   | agtggtccga  | 201300 |
| ggaccaggag  | acagggcagg  | ctaacagttc  | tgcttaggca  | gtagcactgt  | gggcccagaa  | 201360 |
| gacaggtttg  | atggctggta  | tcacgacctc  | acctattggc  | ctcctacatg  | gcaaaggctg  | 201420 |
| ggccagggtcc | tgatgggtct  | tcctctttga  | ggcttctgct  | cctgggtatcc | agatatagct  | 201480 |
| aagtctgtct  | ctgagaagtg  | cacctctgcc  | acagcaccag  | gcctgtaact  | cccttggttg  | 201540 |
| taatagactc  | tatcaactct  | ctcctttcct  | gtctttgatg  | aagtaagcta  | ccactactga  | 201600 |
| gatggctatg  | tgtcaaggaa  | atgagagtgg  | ctattggcca  | acaatgagta  | aaacactgaa  | 201660 |
| gctttcagtt  | tgacagccca  | ctagggaagta | aattcagtc   | gcaaccaagt  | aagcctggaa  | 201720 |
| gcctggatcc  | ttcctcagtt  | gtgccttcag  | atgagactgc  | agttctattg  | gcctttgtct  | 201780 |
| tgacgccttg  | taggaaacat  | tgaagtggag  | gacccagcca  | agcctggcct  | ggactcttga  | 201840 |
| cccacagaaa  | atgtgagata  | taataataag  | tatttggtgt  | tctaagctgc  | tacattgggtg | 201900 |
| ataatttgct  | acacagtaag  | agataatgta  | tacaggttct  | tttcatttct  | tttgctatta  | 201960 |
| caaagagtgt  | ttctgtgggc  | caggcatgggt | ggctcacgcc  | tgtaatccca  | gcacttttgg  | 202020 |
| aggccaaggc  | aggtggatca  | cttgagggtca | ggagttcgag  | gccagcctag  | ccaacatggt  | 202080 |
| gaaacctgtc  | tctacaaaaa  | atataaaaaa  | ttagccagat  | gtggtggcat  | gcacctgtaa  | 202140 |
| tctcacatac  | ttgagaggct  | caggcaggag  | aatcgcttga  | actcaggagg  | tggagggttc  | 202200 |

|            |             |             |            |            |             |        |
|------------|-------------|-------------|------------|------------|-------------|--------|
| agtgagccga | gatcgtgcaa  | ttgcactcca  | gcctggacga | cagagcaaga | ctccatttta  | 202260 |
| aaaaaaaaaa | aaccaaacia  | aaagagtgtt  | gctatgaacg | ttgttgtaac | atgtctccca  | 202320 |
| gaacacagaa | cacaagagcc  | taggtttccc  | ttgggtatgt | gctgggtata | tgccataggaa | 202380 |
| tggaattgtt | agatctcagg  | atgtttgttt  | gtttttgaga | cagagtgtcc | ttctgttgcc  | 202440 |
| taggctggag | tgcagtgggtg | cgatcttggc  | tcactgcaac | ctctgcccc  | cagcttcaag  | 202500 |
| caattctcct | gcctcagctt  | cccaagtagc  | tgggactaca | ggtgcccgc  | accacgcctg  | 202560 |
| gctaattgtt | cgtattttta  | atagagatgg  | gttttcacca | ttggccaggc | tggtcttgaa  | 202620 |
| ctcctggcct | caagtgaccc  | acctgccttg  | gcctcccaaa | gtggtgggat | tacaggcctg  | 202680 |
| agccaccatg | cccagctgga  | tctcaggatt  | tacaaatgtt | ccacattact | ggtagtcaag  | 202740 |
| aaaccattaa | agttagggtc  | atgccatggc  | tgtagtgtgt | tggccagggt | tgggagtaaa  | 202800 |
| gcctggcctc | aatccatgga  | agttgggatt  | gtctgtgggc | aggggtggag | tcgtacagag  | 202860 |
| ggccgagtta | ggggtcagtc  | cttacgtagg  | gttgagggtg | agtgtgtggc | cacatttggtg | 202920 |
| aacaatttga | agtctgggtt  | gggggtgggt  | ttggacaagg | ttggagaaca | ctgtgagggg  | 202980 |
| gggtttgggg | gtcaatctaa  | ggccagggtt  | agaagttagg | ctgtcatcct | gtgcaaaacc  | 203040 |
| tgacccctgg | tgatatcatc  | aacctaccaa  | gctgtggccg | cacaggaccc | agccaccac   | 203100 |
| aagatgagat | ccactctggg  | tccagaaagc  | tctttcactc | aaggcggggg | tggggagatg  | 203160 |
| gaggacaagt | gaagaagaaa  | ctggctcccc  | tcagatgcaa | aaagagaagg | caggcacagg  | 203220 |
| aaatagaaca | caacactgac  | tttaatgggg  | cagccctgag | ccgtaacctt | caaaccttcc  | 203280 |
| gcctcagtag | cccgtgcgcg  | aggagggagg  | ggcgactgct | acgggcacat | cgctgatgtc  | 203340 |
| ctccctactc | ctggcgatcc  | cacgcctcac  | cccttctctc | ccagctgctg | cggctcccg   | 203400 |
| cgccgaggtg | ggtcccagg   | aagctcccag  | gaggggcagg | ggagccctc  | agcgccctcg  | 203460 |
| gacagattcc | cccatggctg  | agggcatggg  | gaactagcct | ggatggagac | gcccgcctcc  | 203520 |
| tccgagctgg | gcggggacct  | gattctgggg  | gtgtggacag | agccaaggga | ccgcccccca  | 203580 |
| aggcccagcg | ccgggagatg  | caaggccggg  | cccaagggtg | tggacactgc | tttgggggac  | 203640 |
| aggctaggtc | tctgcacgtg  | gcttctgggc  | tctggaaagc | ggtccattct | cctgacccgg  | 203700 |
| atctccggag | tggtaggagg  | cggctcagtc  | ccgggcctgc | gctcctagag | ttcctgtccc  | 203760 |
| atctcctccc | acgctcacc   | atcccaagga  | aggagggcac | tcgggcccc  | gcaggctcgt  | 203820 |
| gagagcagcg | ggctccgccc  | tcccaatggt  | ctatccatcg | gtgggtgggt | ccggcgcgcc  | 203880 |
| gtcggggctc | tggcgggtag  | ccgggcgtcc  | ccgcgcggcc | cgggcgcgcg | ctcaccgctg  | 203940 |
| cttacgctcc | gcctgctgga  | gcgcgcggga  | gaggtcttcg | atccgcacat | tcttgagctg  | 204000 |
| gagcaggtgc | tccaggcgcc  | ttaccttcat  | ctgcaggacc | taagcgcgac | cgcgcagcgc  | 204060 |
| tcaagcctgg | cggctctgct  | cctcctgccc  | ggcctctgct | ggcgcgcgag | ccccaccgc   | 204120 |
| agccttcata | gacccgtcac  | ttgcacggtc  | tcttgagagg | ccaacagctc | ctgctccttt  | 204180 |
| tdagcगतct  | ggaggacgaa  | gctggggctg  | ccctgcaacg | cccgggtata | cgctggaggc  | 204240 |
| cgaggcgccg | gcgcccggtc  | ccagctgggg  | tgggaagcag | cttagcggag | gcttggacct  | 204300 |
| cgggccttac | actcaccctt  | ccagggcctt  | agaccaggcc | ttcccttctc | ccctcttagc  | 204360 |
| ccctttcctg | cctcccgctt  | caggctcagc  | gtaccgtgcc | ccccaacccc | agctgtgtgc  | 204420 |
| agagatcacc | agcctgggtg  | agtctggaac  | tcccagcttc | tttacctgag | ctgacccctt  | 204480 |
| cccggggggt | ccgggacacc  | ttcacctcgg  | gccttctggg | atacacctga | gacaaggcaa  | 204540 |
| gtggaggaat | gactgtgctg  | gagctctcatt | ctcctgcttc | ccctcccta  | cttccactct  | 204600 |
| cacaagtga  | aggaagaagg  | acagaaggaa  | agctccaacc | gggagggaga | aatggaaaag  | 204660 |
| gccaccttac | ccacatccat  | gtagccactg  | ccatcctggg | gagccagctc | ctgaaagaac  | 204720 |
| cccggggggt | cgtctgaagc  | ccgggctcag  | tcaccccaaa | ggagccccat | gtcccttcca  | 204780 |
| cgggccccca | ggccaggctc  | tattagccaa  | gcacccctcc | cccacactta | tcacccaggc  | 204840 |
| tggcctttct | gtggtctggg  | gagcgccccc  | caccagcccc | ccagagcagt | gggagaagct  | 204900 |
| ggagcggggg | ggcacctgta  | aggagccggc  | gcctgcttcc | ctgcgcctct | gcctctcctc  | 204960 |
| caggcgctgc | ctcagcggga  | tgagcaccag  | ctccaccacg | cctggggcgc | actgcgcgat  | 205020 |
| cttgcgcatc | acgtcatccg  | gtactgaaaa  | gttcagcctc | ttcagtacct | tcttgagggg  | 205080 |
| tagacattgg | gatagcagat  | tgggtcctgg  | ggaaatgaag | cagggggatg | ggggagaggg  | 205140 |
| aagggggcat | ggtgggttgc  | gttctgagtg  | tacagccctg | ggggagtgaa | agcgcatcgc  | 205200 |
| atctgagctt | ccttggtgtg  | gtgttccgag  | gttcccttgc | tccttcctaa | tagaagagtc  | 205260 |
| tcttgatttt | gtaatcggac  | gaattctgtc  | tgcgttgacc | ctccccaccc | tgccaccaa   | 205320 |
| acagtctcag | cctgggcaac  | caagaccata  | ggagatgatg | gggctgcaca | ggtattctgc  | 205380 |
| tacctgttca | gatgacccca  | gttgctgagc  | ttctgctgga | gagagttggc | ggggacataa  | 205440 |
| ttgtgcatct | ccaccatctt  | ggggaagtaa  | aacttgatga | cctctgcaac | aaggactgag  | 205500 |
| ggagagggga | gcagacatgg  | aaagtggggg  | catgggtaag | gaaggaggag | gtgagagaag  | 205560 |
| gtgggggtgg | gagaatcagg  | catggcagag  | tgggtgaacc | caatggaggg | gtggagatgc  | 205620 |
| ggaaagggga | aggataagga  | gggagacaaa  | gacattaaga | ttagtgatag | gtagaggaga  | 205680 |

66/154

|             |             |             |             |            |             |        |
|-------------|-------------|-------------|-------------|------------|-------------|--------|
| ctaagtctcc  | agcatacctc  | agaggagctg  | catcccttcc  | ccactaattt | aggaatggga  | 205740 |
| atactgagtc  | tctgggtcctc | cagtagagaa  | tgagtcagcc  | agtgcagggg | ctcacacctg  | 205800 |
| taatcccaac  | actttggggag | gctgaggcgg  | gtagatcacc  | tgaggccagg | agttcgagac  | 205860 |
| cagcctggcc  | aacatggtga  | aaccccatct  | ctactaaaaa  | tacaaaagtt | agccaggtgt  | 205920 |
| ggtggtgcat  | gtaaaccag   | ctacttgggt  | ggctgaggca  | gaagaatcgc | ttgaactcag  | 205980 |
| gaggtggagg  | ttgcagttag  | ctgagatcat  | gccactgcac  | tccagcctgg | gcgacaagag  | 206040 |
| tgagactctg  | tctcaaaaaa  | aaaaaaaaaa  | aaaaaagaga  | atgagtctcc | aagagaagct  | 206100 |
| ttcaggagcc  | caggggaattc | acatggatgc  | atgatgcaca  | cacacacaca | cacacacaca  | 206160 |
| cacacacaca  | ctccaacacc  | aaagttccag  | gagagtaaaa  | gactcacgta | agcagagaca  | 206220 |
| catgtgataa  | gatcccccta  | cacatagata  | catagggcag  | aatgtgtaga | cacatagaac  | 206280 |
| aaacacatgc  | tagacacaca  | taggcacaca  | tgtagaacac  | acatagacag | aagcacaaaa  | 206340 |
| agccacatat  | gcacacagcc  | tccccaaata  | tacagatata  | gagatgcata | gaagtacttg  | 206400 |
| acatgaaaag  | atgacagata  | agtattttctg | tcattgagat  | cccaggagac | attcctagaa  | 206460 |
| acacaaacac  | cttcacaatc  | acaaagaaac  | acatgcacac  | aagaccccc  | caaatacaag  | 206520 |
| actaaacatg  | tcaacacata  | tgaacacaca  | catgctcaga  | ctcccagaag | acatagaaca  | 206580 |
| catagcaatg  | agtgagcaca  | tacagaaatg  | cataaatata  | tatgcagaca | catatcttaa  | 206640 |
| tcattaatgt  | ccctggaggg  | agttagaggc  | accccccca   | acatgcacac | aggatagaac  | 206700 |
| atgtagacat  | acatatatgg  | acacacatcg  | ttcacacaca  | gatgcaaata | gaaaaacaagc | 206760 |
| acacacatct  | aaaattttgaa | gtcccaggag  | gcagattcct  | ccccactccc | acaagcacac  | 206820 |
| aaacacaaag  | agttttgcaca | gagaccctgt  | acacacgaac  | atacaggaca | gaacatgcag  | 206880 |
| acacacaagg  | gcacacacct  | ccatcgctaa  | agtcccgga   | gaggtttcgc | ttgggcccggg | 206940 |
| acagagggat  | gttgtctacc  | cacagggtaca | gctggtgcag  | cgctcctcgc | tccacgctgc  | 207000 |
| tcgccattgg  | cgtcctcagc  | gcctggccgc  | cccagcgggtg | ccgggtcccg | ccccagcctc  | 207060 |
| acgacccttc  | aggcgcttcc  | ttctctgcca  | ctgcagaagc  | ccataactgc | ctctgectgc  | 207120 |
| ctaateccaga | gactcacgtc  | ccagctggga  | gcggccatat  | tgttttctga | aaccatggaa  | 207180 |
| accattccccg | actctcctgg  | gtacaaagac  | ttctgggagt  | tgtagtttta | tccccctccc  | 207240 |
| gctacctcct  | ctcaggcccg  | agctttgaaa  | aggcggttag  | ctgccccttg | cttcttccca  | 207300 |
| agaaagcatc  | acttctgtgc  | ccacgccacc  | tagtgaccag  | ccactcatcc | atttttgggac | 207360 |
| cagccaaagc  | caaccaggac  | cccagaattc  | cgcgatcct   | tctatagtgt | cacctaaatg  | 207420 |
| tcgacggcca  | ggc         |             |             |            |             | 207433 |

<210> 6  
 <211> 23574  
 <212> DNA  
 <213> Homo sapiens

<400> 6

|             |            |            |             |             |             |      |
|-------------|------------|------------|-------------|-------------|-------------|------|
| gctggacott  | caacagagaa | ggatggctgg | cagagtccta  | gtctgatagg  | cgcccccttc  | 60   |
| ctagtgcctc  | tggagcagga | gcacagtga  | aaccagctg   | tctgogctga  | tacccttggg  | 120  |
| aggagccttt  | gctgcagcaa | ctcccgcgcc | tcattaagtc  | tccaccccag  | gctgcaccgt  | 180  |
| caggtaaacc  | tgggagccac | gggatctggc | gcctgaatgt  | gcctagacgg  | accccaacac  | 240  |
| actagccctg  | ccacgcacac | ccagcaacac | acagacaacc  | acccagggag  | acctgggtctc | 300  |
| tcccatgcta  | aactaggaac | gctatgaaac | agcatttttc  | ttttcctttt  | ttatataatt  | 360  |
| ctattttaaa  | cttaaaaaaa | tttttttgaa | tagagttggg  | gtctcgctct  | attgcgcagg  | 420  |
| ctggctcttga | actcctgggc | tcaaaccatc | ctcccacctc  | ggcctctcaa  | agtgcctagg  | 480  |
| tgacagggtgt | gagccaccat | gcctggccta | ttttaaactt  | cttattgttg  | actaatttta  | 540  |
| gacttgcaga  | aaagttggaa | aatgtaagca | ttttccttac  | accttccacc  | cagcttcccc  | 600  |
| tgatggtoca  | tcttacttaa | tcatagtcca | atcattacag  | ctaggaaact  | aaccttggta  | 660  |
| caatcctggt  | aacgaaactg | tagactgttt | tgcattttctc | atttttccac  | taatgtcctt  | 720  |
| tttctgttcc  | aagatccatt | ctaggatccc | acatcacagt  | tgtctcctta  | ttctcctcaa  | 780  |
| tctgggagtt  | ccttcatctt | tccttatctt | tacccttgac  | actttttgaag | aatcctggcc  | 840  |
| agttatttttg | cagaatgttt | ctcttgagtt | gtctgatgtt  | ttattctgat  | cagaatgaga  | 900  |
| cacagcattg  | ttttgactaa | ccaaaaagtt | attctataag  | taaataattga | ggtaaaaaat  | 960  |
| ctcatccaac  | ctgggcaaca | gagtaaggcc | ctgtctcaaa  | taaagtctca  | acactaagat  | 1020 |

|             |            |            |             |             |             |      |
|-------------|------------|------------|-------------|-------------|-------------|------|
| ttaaaaagtg  | accagaaaag | ccccactat  | gatttgtctt  | gacttttttt  | taaaaaacia  | 1080 |
| acaaacaaac  | aaacaaacia | acaaaaacga | tgtcttgcct  | tctctctccc  | tcaggctgga  | 1140 |
| gtgcagtggc  | acgatcttgg | ctcactgcaa | cctccgcctc  | ccagggtcaa  | gcgattcttc  | 1200 |
| tgcctcagcc  | tcccaggtag | ctgggattac | agacgcccgc  | caccatgccc  | ggctaatttt  | 1260 |
| tgtattttta  | gtagagacaa | ggtttcacca | tgttggccag  | gctgggtctg  | aacttctaac  | 1320 |
| ctcaagtgac  | ccgcccacct | ctgcctccca | aagtgtctgga | attacagacg  | tgagcaactg  | 1380 |
| cgcttggcct  | cattagcatc | ttaaatctcc | acacaggggt  | gtgttcctta  | ctgttataag  | 1440 |
| gagcaaagga  | tcagtttgag | gacaggtaaa | ataaaaatgc  | gcttgcctgc  | tagagggaga  | 1500 |
| agtccttgc   | gaagatagct | ttgcttgaat | gagctcaatt  | gcaatgccag  | tgctgaggct  | 1560 |
| tgttgactgt  | acggtcacca | cagttgctgc | tgcgcgccta  | gaacatggtc  | actttcttga  | 1620 |
| ctacctatcc  | tgtctcagta | catctgtctg | tggtttgttg  | tggtccattt  | cctaattttt  | 1680 |
| ttaatgaatc  | agaagactgt | gatgtgcttt | ccgctgtgct  | aacctaggcc  | gctgaagcaa  | 1740 |
| aatgtaaacc  | aagatgcccc | tgcagtgggt | gtgcttcact  | ctacgacatc  | tgttaccgga  | 1800 |
| aaggggtcca  | gattcagacc | ccaggagagg | gttcttggat  | ctcgtgcaag  | aaagaatttg  | 1860 |
| agacgagtc   | ataaagtga  | agcacattta | ttaagaaagt  | aaaagaataa  | agaatggct   | 1920 |
| actccataga  | gagcgcagcc | ctgagggtcg | ctgggtgccc  | atttttatgg  | ttgtttcttg  | 1980 |
| atgatctgct  | aaacaggggt | ggattgttca | tgtctccctt  | tttttagacca | tatatggtaa  | 2040 |
| cttctgata   | ctgccatggc | atctgtaacc | tgtcatgggt  | ctgggtgggag | tgtagcagtg  | 2100 |
| gggaccgacc  | agaggtcact | ctcatcacca | tcttgggttg  | ggtgggtttt  | agccagcttc  | 2160 |
| tatatgtcaa  | gctgattttt | ttgggtgggt | tggtttttga  | gacggagtct  | cgctccaggt  | 2220 |
| tggctgtcag  | tgacatgatc | tcagctcact | gcaacctcca  | cctcctgggt  | tcaagcgatt  | 2280 |
| ctcctgcctc  | agcctcccaa | gtagctggga | ttacaggcac  | acaccaccat  | gcctggctag  | 2340 |
| tttttgtatt  | tttagtagag | atgagggttt | accttgttgg  | ccaggctggg  | ctcgaacttc  | 2400 |
| tgacctcagg  | tgatccgccc | acctcagcct | cccaaagtgc  | tgggattaca  | ggcgtgagct  | 2460 |
| accgcgcctg  | gcctactgca | aactatttta | tcagcaagg   | ctttatgacc  | tgatatctct  | 2520 |
| atctcatcct  | gtgacgcaga | atgctgtaac | tgtctggaaa  | cgcagcccag  | taggtctcag  | 2580 |
| ccttattttt  | ctcagcccct | attcaggatg | gagttgctct  | ggttcacagg  | cctctgacac  | 2640 |
| atcctcttgt  | gttttggcgt | gtgggaggaa | agaggggtga  | gggaaggaaa  | ctcaaaacca  | 2700 |
| agctctgacc  | acacagggca | ggtacactct | cccactgtc   | tgtgggtgcc  | acaagtcaag  | 2760 |
| ggagggggcag | agagagaaga | aggtgtgaca | gatggccgca  | ggccacagaa  | tgtagagga   | 2820 |
| agcccagttc  | ctcccggggc | agcccaagta | gctggtagtt  | gggtggccaa  | acagaggcgc  | 2880 |
| tcacagctga  | gctgggctcg | ctcgctaccc | ccagctcagc  | gtccactctg  | cccctcagta  | 2940 |
| cctcctgtct  | agcctcaggg | tcctatgcct | cctcctgct   | tcccagtcac  | ttctgcgtgc  | 3000 |
| ctcctgcttt  | tctgctgttg | gccccatgcc | agctcctttc  | tgttgagctt  | ctcttctcca  | 3060 |
| gttctgcagc  | acagccaggt | gatcctgggc | tccagacagg  | cctctccccc  | agtctggggc  | 3120 |
| ctccccctct  | agagccctct | ttccttccca | cgtggcctcc  | ccagggttcg  | ccactgaatg  | 3180 |
| gagaaggggt  | ggaggggggt | ctgggcagtc | ttggggatta  | gccaagaggg  | cagagttggc  | 3240 |
| ctccccaggg  | tcccttgtag | ctggagtccc | gcggggctca  | gacaccccc   | cctgaagggt  | 3300 |
| aagagcgggg  | agggtatatt | aacgtgtatt | tttagagtct  | ctcttttttt  | tttttttttt  | 3360 |
| tttgagacgg  | agtcttgctc | ttgttgccca | ggctggagtg  | cagtgggcag  | atctcagctc  | 3420 |
| actgcaacct  | ccacctccca | tgttcaaaca | attctcctgc  | cgcagcctcc  | cgagtagctg  | 3480 |
| ggattatagg  | cacgtgccac | cacaccgggc | taatttttgt  | attttttagta | gagatggggg  | 3540 |
| ttcaccacat  | tggccaggct | ggtctcgaac | ttctgacctc  | aaatgatcct  | ccgccttgg   | 3600 |
| cctcccaaag  | tgctgggatt | acaggcgtga | gccaccatgc  | ctggccttag  | attctctatt  | 3660 |
| ttatgatggg  | ttaacatctc | ggggtggggg | cttgttggct  | ggagagaaac  | tgcttgattc  | 3720 |
| ctggagatca  | gaaacaaact | atgcctttca | tatgcaaacc  | gaccagtctt  | gagtccatac  | 3780 |
| accaaccacc  | cccttcaagg | aactctcaca | tacgaaacca  | gtatttcccc  | tgccctaaac  | 3840 |
| cagctcaggg  | caggcacct  | accagacaa  | ttagagccca  | acccacgcgc  | ccaaacccca  | 3900 |
| ccagaatgat  | tcaaaatgcc | aatcctaccc | tcttccccct  | gcctgccttg  | cctccccagt  | 3960 |
| ggaaacggca  | ctgtgggctg | tggcctgtgc | cttccactcg  | ctcctgattc  | tgctccctgga | 4020 |
| ccaaacctag  | tgccctccca | ctgtggccct | gcatgggtgg  | aaactgtgag  | taataactta  | 4080 |
| tttcaacagc  | attggcctct | gtgtcatcag | tcaccttcac  | aaattaaaa   | tttgcaacta  | 4140 |
| caactgaggg  | aggagggggc | tttagagagc | actctcacct  | ggcccccttg  | gcttgtggag  | 4200 |
| tggagcccga  | ggtgaagtgt | cctccctgca | ctcagttttg  | ggatgggttt  | gttatgcggc  | 4260 |
| aacagctggc  | tgatataggc | cactcagcag | ctcttgcacg  | aagcaatggg  | agaatgtgaa  | 4320 |
| cgcccaaggg  | aggcaggagt | gacagagcaa | agaggggtgt  | caaactaggg  | aacaccgtcc  | 4380 |
| tgtgccccaga | cacatgcctg | gggctctggc | tgccatctag  | gctcggcctg  | ccagccacca  | 4440 |
| ccccgtcagc  | caccacccca | ccaaccacca | ccccgccaac  | caccaccccg  | tgtggctcct  | 4500 |

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| agggcacccc  | atggggtgaa  | tttacataca  | gaaaagacta  | accaagggtcc | aagttaatat  | 4560 |
| gtcttttttaa | aattttatttt | agagacaggg  | tcttgctctg  | tcacccaggc  | tggagtgacag | 4620 |
| tgggtgccacc | atagctcact  | gcagtgttga  | actcaggctc  | aagtgatcct  | cctgcttcag  | 4680 |
| cctcctgagc  | agctaggcct  | acaggtacac  | accaccacgc  | ccagctaatt  | ttaaaatttt  | 4740 |
| tctgtagatg  | cgggtgtcttg | ctgtgttgcc  | caggctgggtg | tggacctgct  | ggcctctggt  | 4800 |
| gacccctcta  | ccccggcttc  | ccaaagtgtc  | gggattatag  | gcatgagcca  | ccacgcctat  | 4860 |
| agccaacatg  | tctttctttt  | gacttctact  | ttggtatctt  | ttcttaaattg | gttccctctg  | 4920 |
| tccccccgac  | acacacagaa  | tgggggagag  | gctgtcagat  | tctgagctcc  | agaacctcag  | 4980 |
| gtgtagcact  | gggattgggg  | gtgggggctc  | aggaaccacc  | taggggagaa  | gacagggtgg  | 5040 |
| gaagaaacag  | gaagggaagt  | cccaaaaatt  | atgtttgttt  | gcagaggcca  | gccaggctgc  | 5100 |
| aggggagtgt  | ggactcagtc  | gaaccatagg  | gccccaggac  | cactagcttc  | tggccagcag  | 5160 |
| tcatgccttc  | cacagagctg  | ggtccgtgga  | aattgcatgt  | aggagacaca  | ccagactccc  | 5220 |
| aggacagagc  | ccttttgagg  | tggccagcac  | taccagcct   | ccactggtga  | gggaggtcag  | 5280 |
| gggctgtgtg  | acctttgctt  | ctgggactga  | tggtttattg  | agctggagag  | tgtgccagc   | 5340 |
| agtgttctcc  | agccctcagg  | aacttctaatt | gtggctctgg  | gttcctggag  | tgggtgggtc  | 5400 |
| gaagctccac  | tgggggaaga  | aacttccaag  | ctgcctgcag  | gtgctggagg  | tccggtgatt  | 5460 |
| cactggctct  | gccccctgcag | ttcaagttcc  | tggagtggct  | gtcagtggcc  | acctgtcttt  | 5520 |
| aaatctgttc  | attttaggag  | ctacctctca  | ccagaggcag  | gatcttgga   | tctggacttg  | 5580 |
| atctgtcag   | aatgaggagg  | atatgttgct  | ccctaaggac  | tggggcccca  | ggctgcaagc  | 5640 |
| tgtgtggcag  | agagcccatc  | ctcactcagt  | gaggaccagt  | gatccaggaa  | aagccacagc  | 5700 |
| ttctccctcc  | ccagccagg   | ggcttccagc  | atcctggctc  | ccatgataac  | caagagggtca | 5760 |
| taaactcatt  | tcataataa   | cctgagccca  | gaaacctgat  | tagggggcag  | caaactgagg  | 5820 |
| ggtgggagag  | gtgggagggt  | gggcgatgag  | agggggaggc  | tttgaatcca  | ggtccctgcc  | 5880 |
| taccttgggg  | gtcaggcgag  | actgctggca  | gaggcttctc  | aggggtggctg | ctgggctcat  | 5940 |
| gagagttctc  | aggggtctggg | agaaatgggtg | gagggtaaat  | gttgtgaata  | tggtcagcag  | 6000 |
| gagaccctgg  | ggctggggag  | gggcataggg  | gactcaagggt | gactgggtgc  | tgcccatctg  | 6060 |
| gaaggaggga  | ggaggcatga  | gcccttccct  | tctcccttcc  | ctctccacct  | ccccctggtg  | 6120 |
| cctcactcac  | ccaggggcca  | gggctgtcca  | tgggctgtgg  | ggcccaactc  | catggggtag  | 6180 |
| acgccgccca  | gggggtggctc | cctgtgtggg  | ccatctttgg  | ggctgagcaa  | cgtgataaga  | 6240 |
| gtccaggagg  | ttggcacagt  | gatcctgagt  | gggttattgc  | ctccccgcag  | catggtgtcc  | 6300 |
| agcccaggga  | gttctgcgtt  | tactgagttt  | cttggggcac  | ccatctgctc  | caagtcaacc  | 6360 |
| tctcagctcc  | cttctgctc   | cctcttcagg  | ggagccttgg  | gatccaggct  | ccaagtgagc  | 6420 |
| ctcatgcect  | cggctggcac  | ctcctctctc  | tagtcctaac  | atcttctcca  | ggctctgaca  | 6480 |
| ccaccagca   | gcctggcact  | ctccagatgc  | tggcatcgct  | cagcttccaa  | agaaccttgg  | 6540 |
| atgtccgccc  | cttcggcagc  | tatgtctgct  | ctccttgccc  | ctgggtgccc  | tgttgccttt  | 6600 |
| gatgattcca  | agccatcttt  | gactgtcccc  | atcccatccc  | ccaaggcctt  | gtcatttctc  | 6660 |
| gtgatgttcc  | ttcaaaacat  | tctcccttgc  | cctgagactc  | ccgcctgggg  | atgagaagca  | 6720 |
| gccgccaccc  | tctgcagcgc  | cccctccgtg  | ctgacacgcc  | aggctctggc  | caccttgctc  | 6780 |
| ctctgcccac  | agacctcaa   | tcacaactcg  | ctttgtcagg  | gtcctcttag  | ctgccacccg  | 6840 |
| ggggccagggt | ggtgccctgc  | ccctgtctgt  | tatgccctct  | gcccccatct  | ctggcccaaa  | 6900 |
| tcatgccatc  | tcccttggct  | tggccggagc  | actcccaaga  | ccaggctatg  | tcagacatgg  | 6960 |
| ccacagagtg  | cctgccctgc  | ctagggccct  | ggtgcagggt  | gagtcctagg  | acagccatgc  | 7020 |
| ttagtattat  | gtgactcccc  | actccgccac  | caccaggtc   | acagagaact  | gggttaaggc  | 7080 |
| agggccctgg  | cacaggggca  | gccagcaccg  | cagctgacca  | gtggtatgga  | gtgaaaagat  | 7140 |
| gtgctggggc  | cagcatttgg  | gaacttcaag  | ggggtgacag  | aggtgatttg  | tgcagaggaa  | 7200 |
| gtggcaaagg  | gocgaaaact  | ggtgagacag  | aggctggaca  | ggcctccggg  | ggcagcatgg  | 7260 |
| tacagggact  | gcaatctgag  | ccagggaaaa  | acagggcgaa  | gtcaagggtg  | aggcagccag  | 7320 |
| ctgggtgggag | aagcaggaga  | gtggacaaga  | ggagctgtac  | tgggaggtag  | agggccatgc  | 7380 |
| cttgcgggtgc | tgggtggggg  | gcagggatcc  | acccccctgc  | ttgactggga  | ggccactgga  | 7440 |
| acctcctgtt  | caaagetact  | tctttccatg  | gcctctgggg  | ctgctctctg  | cacctggggg  | 7500 |
| caaggctgag  | ggcctgcccc  | agctcccaca  | gccccagcag  | agctctgagg  | aggggaaccg  | 7560 |
| caggagtagg  | ctcaggaagc  | aggcgctcgg  | agcctaccca  | ctgcacgcag  | ggctccctct  | 7620 |
| gcagccccag  | ctgcctcgct  | gcagatgggc  | tcctgggagt  | cggtagcaac  | accaggccag  | 7680 |
| gccggccctt  | gggagcagag  | gcagcaggac  | gctgaggagc  | atggccagca  | ggaagggtgc  | 7740 |
| atggtctgcg  | gggattgggg  | gaagggggcg  | tgagtcctga  | gcagggtgcac | caccccagct  | 7800 |
| cctgccccac  | tgcccccac   | tggcatactt  | tcagcttgca  | cagggccact  | gtccatgctg  | 7860 |
| ccaccaaaagc | ctggcttgct  | acagaagggt  | tgagccagc   | ctggagcaca  | gtggcagtta  | 7920 |
| tggttgctat  | tgcacacctg  | cagagaagag  | aagaggaggg  | tcacgtagga  | ttaggaaccc  | 7980 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| caaggtcacc  | cccactcctc  | gggctctcac  | cccgtggctg  | tggcagggcag | tcaggcagcg  | 8040  |
| ctgaagctcc  | tgggaaggcat | tcttcctgca  | gcgcctgctc  | tggcacacct  | acgggcagtg  | 8100  |
| caccagggcag | tgaggggggac | actggcctgc  | gggattcaaa  | cggcaaggag  | gggtcgggtg  | 8160  |
| ggcagagctc  | accattctag  | gtccacactg  | ggtagcctggc | tctaccaggc  | ccaggccaag  | 8220  |
| cagggtccagc | tgggcactgg  | ggagtgccaa  | ggctccccga  | caagtcactt  | cctggccatc  | 8280  |
| taggtgaacg  | gtagagtcca  | ctggcaccat  | gtgcgggtgcg | agcaggctgg  | gctttccacc  | 8340  |
| ctggcactgc  | agcttcccac  | acagggcatc  | cctggggagg  | aagtagaggg  | gggtcaacag  | 8400  |
| ctgcagtacc  | ccccttcccc  | aaacccactc  | catagcttct  | gctccctcca  | ctcagctcca  | 8460  |
| ctccctacct  | ccctgcacag  | ggcaggaagt  | ggccctcgct  | gtcctggccg  | cagtttccat  | 8520  |
| gagcatctcc  | cgcagagttc  | accacctgga  | aacaggcctc  | gggagctggg  | tgggagcctg  | 8580  |
| agggaagcatg | ggccaggctg  | ggggcagctc  | ggagaggggc  | tgcgctcagc  | gggctcagt   | 8640  |
| ttccctgccc  | catcttcccc  | acagttagaaa | actggcccca  | ccagaggatg  | gggggcaggg  | 8700  |
| tgcaaggggtg | ctcgtgtcct  | ctcaccaggc  | ccccagagct  | gctggcactg  | ctgctccagc  | 8760  |
| gtgggacatg  | cgccatccca  | gcagttagcca | ctgcccctgg  | cacaggggtga | gccgtccagt  | 8820  |
| aggtaaacgt  | ctggggggaca | gtgggaggag  | gtgcccgtgc  | aaaactcagg  | gaggtcacag  | 8880  |
| tcacccatgg  | cctggcgcca  | cagcgctcca  | gccggcttca  | gctgcgcagg  | tgacgggtgg  | 8940  |
| tggggaaaggc | agagagaggc  | cacgtgcagt  | gagaggtcca  | tgccgagagc  | gcggctcggg  | 9000  |
| gctgggggag  | ccaggcctac  | ccaagcccag  | cacccaacgg  | gggaacctga  | gggcaccaat  | 9060  |
| taactaaaggc | caacaggccg  | gctcccaagc  | tccccgaaac  | cctcaccctg  | aaccttccat  | 9120  |
| gcccctacac  | ggcagcgcac  | gcagcagctc  | ccgtggggcg  | actgggcccc  | cgggcgcagc  | 9180  |
| gcagcattgt  | gagcaaaagca | gcagaggtcg  | cggcactcct  | gggaccagaa  | aggcaagaaag | 9240  |
| ggcccagggtg | agggcgcgagc | gccccagacc  | tgagcggaga  | gggcaagtgg  | gggccgggctg | 9300  |
| agccgactta  | acctggccag  | ggccgcagtc  | acactcctcg  | cccgttcca   | cgaagccgtt  | 9360  |
| cccgcagagc  | gccggcgcca  | ccgggagtc   | ggggctcggg  | gcattggaga  | ggcaagcgcc  | 9420  |
| gcccccttg   | cggaagaagg  | cgcgcagctg  | gcggcggtg   | caggcgctga  | acacgcgcgg  | 9480  |
| aaacgggtgc  | ctaccggcac  | ggggaggcca  | ttgggcatgg  | agggacagtc  | ccccaaacccc | 9540  |
| cgcgcttctc  | tgatcccccac | ccctgggctt  | ggctacagcc  | gccagacggc  | cagagcccag  | 9600  |
| agaggggaag  | taacccgcgc  | aaagtccac   | aacaagcggg  | acaggggacg  | atgcggcccc  | 9660  |
| aatagttagc  | agcccgggac  | ccaaggtgga  | atcgcgaccc  | gacgggtgctc | ctcccgtgtg  | 9720  |
| aggagtaacc  | tgcggaagg   | actcggaaaa  | taatcttcat  | accgttgaga  | atccactttg  | 9780  |
| cctgagcttc  | ttcccttttaa | gcctcataaa  | ccaccctgaa  | gcggacacta  | tgatcattat  | 9840  |
| ccccatttta  | cagaagagga  | aactgagggg  | cgaccaaaga  | aacgcagcgg  | aggaaagtccc | 9900  |
| caggactagc  | cgccccgcgc  | cagccccgac  | ccccaccgc   | cgtaccgcgg  | ggccgcagcc  | 9960  |
| atgacgcagc  | ctccggactc  | ggccgcagcc  | tccacgcagc  | agccgtcggg  | gtcgtggctg  | 10020 |
| aggccgaggc  | tgtggccgat  | ctcatggggc  | atggtggctg  | cggcgccgat  | ggggagctcc  | 10080 |
| gagtggctct  | ggggggccgt  | gggagggcgg  | tactgcggc   | cgtagagcct  | cctgtctctc  | 10140 |
| cctcgccccc  | gcccgcgggg  | ctcaccgtgc  | tcacgcctcc  | cgagctctcg  | gcgcggcaca  | 10200 |
| tgcctctgac  | gggcgcaggg  | cccactgtgg  | cgccctggaa  | ggcgcgcccc  | ctggggggcgg | 10260 |
| agcgcggcgt  | gaccaggcgg  | ggccgggagg  | tgagggcgcc  | ccacccgga   | cccgcgtccg  | 10320 |
| ggtcagaggc  | acccacgtga  | gcagctgcgc  | ggagtctggg  | ggccgctgcg  | cccacagccc  | 10380 |
| ccggcgccac  | tgcaggaagg  | cccagagcgt  | ggcgttggcg  | tcctgcgtga  | cgcggctgcg  | 10440 |
| gtcccgtctg  | gtccacacct  | ccaggccggg  | cagcgccacc  | tgaatgtcca  | gagtcctgag  | 10500 |
| aagctgaggg  | cgaggcgggg  | ctgaagccgg  | gacagggcgc  | cccatcgcg   | cgggtggtcct | 10560 |
| tctgggggcg  | cccttccctc  | tccccaaacc  | ccaccagcac  | ctgcctgtcc  | tgcgcgcgcc  | 10620 |
| acccccatca  | ccgtctctct  | ccgcgcggcc  | ccaacctggt  | ccacgtagtt  | ggcgacttcc  | 10680 |
| aggagacgct  | gtttggtgtg  | gttcaagttt  | cgggtgccgag | tcaagaactg  | gggaaggcaga | 10740 |
| aatcccgggtg | gcttgagggg  | ctgagctggc  | cccatccctg  | accccgccaa  | cccttggggg  | 10800 |
| ctctcctcac  | caggggtgtg  | tctgccacaa  | tgtacagttc  | caggtacttc  | cgggtcctgc  | 10860 |
| gcgcttctcg  | cctgccctgc  | ggaggtgcaa  | atggggaccc  | tgagtggaaag | ctgctgggct  | 10920 |
| tgagccctga  | cccccaaccc  | cagctcccag  | aaggaaagttt | aacatgtttt  | ctggaaacttg | 10980 |
| tttcttcaga  | cttcaataaa  | aatactggga  | ctcgaggcct  | gtgaattcct  | gtctcttctg  | 11040 |
| atgtggaggg  | ctatagatac  | agcattccca  | ctcccatccg  | atcgatgccc  | ctgacctgc   | 11100 |
| tctggggacc  | accaggaagg  | ctggtcatgc  | ccgcttctgt  | cccaggatcc  | ctgtggccac  | 11160 |
| aggttccttt  | ccaggtgagc  | agctgctcca  | tccgaaagat  | ctcgtgggtt  | gagaagtcc   | 11220 |
| tggagccccg  | gggtggccag  | ggacgcagat  | aatagctggc  | attcctgctg  | aggggtgatca | 11280 |
| ggccactagg  | gtgcagaggg  | gtaggagcgg  | gtgtgagggg  | gctctttccc  | catcccaggc  | 11340 |
| ccagcctcct  | ctcccagagc  | tcacctcatc  | ccagagcagg  | tgcagaggac  | taccagggag  | 11400 |
| tcgggggaagc | cccttactcg  | cccttggtag  | tggcaatgat  | cctagggagg  | aagggggccag | 11460 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccccaaatct  | cagccagggc  | tggagcaaga  | ggggcaagag  | ggaggggtgtg | gtaggggctg  | 11520 |
| gctccaaccg  | ccccttagga  | atgcaaggag  | gagtaggggt  | aggaatgggtg | gggggggtacc | 11580 |
| tctggcggtg  | catcccagag  | cccatggaag  | catctcaccg  | tgtgggtggg  | ggccagcacc  | 11640 |
| actggctgcc  | catctgggcc  | gtagtgggtt  | tctatgtatc  | ctggggccag  | cagcctgctg  | 11700 |
| agaggggggtg | ttacaggga   | caactgaattc | agcttctctc  | tgcctcctcc  | aggatgtctc  | 11760 |
| ccagccttcc  | tccctaaatg  | ctaattggagc | agctttatga  | gtgagacact  | cacagtgtgt  | 11820 |
| cttagggaag  | ggacaggagc  | aatgggtgact | tgctcagatc  | agaaactctt  | ggggctagag  | 11880 |
| gaaggagcct  | tgggtgatggc | ttagttgtgg  | gagatgtgaa  | tatgggaagc  | accagggagg  | 11940 |
| acgccgggga  | ggagtgggaa  | taggggaaga  | gtttgtgggt  | cccaggggac  | ctgcagcagg  | 12000 |
| cagcaggatc  | cacaggatcg  | ggaggggagg  | agtccaggaga | caactgccga  | gaatgggact  | 12060 |
| tggagtggg   | gaaatgcggt  | gacctcccc   | agttccccctg | cctgctgccc  | tcctttgttg  | 12120 |
| ggcatctggt  | cgacctctt   | gccccacct   | gccctagatc  | cttgaaatat  | tttctctaga  | 12180 |
| cttctagacc  | ccacatacct  | cccacctgtc  | cttcagtgat  | tgatgctcac  | cccctgcctc  | 12240 |
| cagagaaaac  | agaatcgcca  | cctgcccacg  | ctgcttccac  | cctccctgcc  | ttctccacac  | 12300 |
| ccactccggt  | aatgattcca  | tcttcaggct  | ccatctcaac  | aggatctttc  | ccacacggat  | 12360 |
| ggatcaatca  | taagtcaatc  | tgtcttcttt  | aaagaaaatc  | cttaacccaa  | cctcaccttg  | 12420 |
| gcctcattac  | ctccagacca  | cccgtcaatg  | atggctgctt  | ccccctccc   | aggcattcca  | 12480 |
| ccacctgccc  | cagctctgcc  | ccctacccct  | gccccacaca  | cacccgccac  | cctaggagggt | 12540 |
| aggtgatgtg  | accaccccat  | tgaagggtga  | gggacgtcgg  | gaaaatatgg  | ttgggcacag  | 12600 |
| tggaaataga  | gtttgttccc  | tgtccatccg  | actccacgag  | ggagaataaa  | atacgtgtca  | 12660 |
| agtgtcaga   | acagcgctg   | cggccaagca  | ctcagtagg   | gataataact  | gataacataa  | 12720 |
| tctgggtggt  | tttaagagcc  | tgcgtccag   | ccggacacc   | cacccacccc  | agcccaaggc  | 12780 |
| accttctctga | gcacaggtct  | gcctgtccct  | tcccagctc   | aaaatctttg  | gctctggatg  | 12840 |
| gtccaggaca  | ctgagcacca  | aaatggcctt  | ctgtgatctg  | gccctcctgg  | gactctccaa  | 12900 |
| attcattccc  | ccctgctccc  | ctccctgtag  | atggagacct  | tccaggcagg  | tcagacaaac  | 12960 |
| tgctgcccc   | aagtgtagcc  | actgcatctt  | tttctttttt  | cttttctttc  | tttttttttt  | 13020 |
| tctttttctt  | ttttttctct  | ttttttgaga  | cggagtctca  | ctctttgccc  | aggccggagt  | 13080 |
| ggtgcagtgg  | tgtgatcttg  | gtcaccaca   | acctctacct  | ccggggttca  | agcgattctc  | 13140 |
| ccgtctcagc  | cccctaagta  | gctgggatta  | caggcgccgc  | caccacgcct  | ggctaatttt  | 13200 |
| ttgtattttt  | agttagagcg  | gggtttcacc  | atgttggcaa  | ggctgggtctc | aaactcctga  | 13260 |
| cctcagggtga | gccgcccgc   | tcggcctccc  | aaagccaccg  | catcttggtc  | cctgccattc  | 13320 |
| ccttagcctg  | gggtgccggc  | tcactttttc  | cctctaggat  | ttcttttagac | tcagcatatc  | 13380 |
| ttgcaaagt   | ccactagggtg | gtgctcactc  | atcgccagca  | gggagctaac  | aagccgctcc  | 13440 |
| tgggggtggg  | agggcgagg   | tgccccacag  | cggggctgac  | agcctcagcg  | gtcctcttca  | 13500 |
| gcctccagg   | agccaaccac  | aggcctgcgt  | gactctccct  | gtcatctgca  | ccctctctgg  | 13560 |
| ggctcctctgc | ccatccagcc  | acccgcacag  | atctgtgtca  | gtccctgccc  | cccaacactg  | 13620 |
| atccccctct  | cccagcccta  | cccagcctg   | gcactcactg  | gttcttctcc  | agctcaagca  | 13680 |
| ggagctcctg  | gccttcagcc  | tccagggcca  | ccagcccat   | gtctggcttc  | gagacctggg  | 13740 |
| caagaaagag  | tgtggagctg  | agatggtggc  | ctccaggcct  | cctgcctgcc  | agggagttag  | 13800 |
| tggcctgtgg  | agccggctgg  | ggaggaagtt  | cttggggaga  | acgtgggctg  | gggagtcagc  | 13860 |
| aggaccccc   | acatactatg  | gagggcgtgg  | aggaggtgag  | aacatacaaa  | gatgttccca  | 13920 |
| aactcaggat  | gtttgcagtc  | ctgacaacag  | ccacttggaa  | gggcgttggc  | acagcctgcc  | 13980 |
| aggcacacca  | gcatactccc  | tagagaccag  | aggtcccaga  | aaggtgcccc  | tcccctggcc  | 14040 |
| ccgccctctt  | ctttcatgcc  | cagaaggggc  | atcaaaagca  | ggggaagaca  | gaggggtgct  | 14100 |
| gaggacatta  | tgggggcatc  | gggtagccat  | ggtcagggcc  | tcctcagagc  | ctctgctacc  | 14160 |
| tgaggcttgg  | ttccaaatga  | gctgctgctc  | atcttctata  | gaattcaaat  | ttgactcctc  | 14220 |
| cacttccaat  | tttggaacac  | tgtccctctc  | tccaaagttt  | tcctgggcct  | ccagcagccc  | 14280 |
| ccgtccctcc  | ggctccgaca  | cctgcttcac  | tggaccacag  | aagtaaacat  | ggacggccag  | 14340 |
| ccagccaaga  | gagcacactg  | gctctcagct  | aggtgtcagg  | aggctggctt  | ggacggccag  | 14400 |
| ccctctctcc  | ttccccacc   | ctcttggcgt  | ctccaccct   | gtgggaacac  | cccacttccc  | 14460 |
| ccttgtccac  | tcagcctggc  | tgggggccc   | gagttggagc  | cggcccagga  | gcttctctgg  | 14520 |
| aggctgctgc  | gccttcggaa  | tgtttaaccc  | ccgactcctt  | ttctccaaaa  | atgcactggc  | 14580 |
| ctggggccct  | gtccaagggt  | ctcagagctc  | ttggaggagg  | ttcttccctc  | gcaagtgggg  | 14640 |
| agcagatgg   | ccttgccctcc | ctggccacag  | gccccacaag  | gcctccagca  | tgagctcatg  | 14700 |
| aggctggaat  | gccacttgct  | ttattgggga  | aaggtctgca  | ccgggaaaaa  | ggccatactc  | 14760 |
| gaggtccctg  | ttcctctgca  | gcccctgcta  | tctttactct  | tgccctcctg  | gtaccctgcc  | 14820 |
| ccttgatata  | taccctcat   | cttgaaatgt  | gagtgtttcc  | tgccttttgg  | aggggatacc  | 14880 |
| tagcctctac  | tctttcttct  | gtaccatctt  | ggcaggcttc  | ctgggggcag  | gggcccaccg  | 14940 |



|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtgggggaag  | cagagccccc  | ttggggctct  | cctcttggtc  | acagcccagg  | ccagacagac  | 15000 |
| agggaggccc  | agaggcagag  | tgaccccagt  | gtgtgtccag  | ccttcccctc  | ctggggatgg  | 15060 |
| ggagggcaat  | ctcaaagctc  | aggccagtgc  | cgtgcttgac  | cagtggaaatg | ggggccttat  | 15120 |
| gggcctaggg  | gatcccagtg  | agggccctgg  | gttgggagct  | gctgggtctc  | tgggggcctc  | 15180 |
| tcagccttca  | tggcaatgct  | cccctgcctt  | ccctcttgct  | ggatttggac  | agtagggctg  | 15240 |
| aaaattccaa  | acaaagaggg  | ctctctagga  | ggggcagggg  | tgtagccaat  | ggtttaaaat  | 15300 |
| cgttcagacc  | ttagtgggtc  | tcaggctccc  | agcctaaaga  | gctgtgtgac  | catggacaat  | 15360 |
| ttccccaaagc | tctctgggct  | tccgtttgcc  | cctctgtaaa  | atgagcatat  | caaggctact  | 15420 |
| gcccctcttag | tttgcagcac  | agatattatg  | gcacaaacag  | atggggcatg  | gttattcttg  | 15480 |
| aagcgtgtga  | agagcgggat  | tgggaagagg  | ctggggcaga  | gcgtcctgca  | gaagaagcac  | 15540 |
| atgggggtgg  | cttacatctg  | ggggacatca  | ggagagtgc   | cactgcccc   | ccataccag   | 15600 |
| aagtggattc  | cacaggagcc  | agtgaggctg  | aaggttcagg  | ccttcgtggc  | agggccctga  | 15660 |
| gagggacagc  | agtgtgtcca  | cagggtcaca  | tgttctggct  | aactttgcaa  | aaggttttct  | 15720 |
| ttttgggtgct | tttttttttt  | tttttttttt  | tagaggctcc  | tgaaaagctt  | caggaccac   | 15780 |
| aaactctgga  | cccatttctg  | cctgggtggg  | gtgggggtgg  | cccagatcat  | ccaggggagg  | 15840 |
| agggaaagag  | ggaggtgggg  | tggagaaagc  | tgaaatgact  | tccatgtgtg  | cgggctcacg  | 15900 |
| agatccagat  | gtccaaaccc  | cagtgccttc  | ttctgcccac  | ttgaggggca  | ggggaggcag  | 15960 |
| gggcctatag  | gagtagtgac  | ttgggtggttc | tggggacccc  | agcaaaacta  | gaagctgtaa  | 16020 |
| tgtagggaga  | gacaaaaggg  | ctgggaggtt  | cagggccctc  | gtggagggcg  | gggagacatg  | 16080 |
| gcatgaccg   | gctcctccag  | gctgacggtg  | cgccagggtt  | gtccatccag  | gaccagtg    | 16140 |
| ggggtgactg  | gctgcccagg  | gatatgtcct  | ggagtaaaga  | cagagcacag  | ggtgaggggg  | 16200 |
| acctgaggaa  | cacaggggca  | tgggacaaag  | cagagggagg  | ggggtagagg  | acatccccag  | 16260 |
| ggagggcactg | gaggcctttt  | ggggcagact  | tcaccttcaa  | cacgcgtggg  | ctcagcctgg  | 16320 |
| agaaggagg   | acgcccgtgg  | gcaccccttg  | atctgaggag  | ctatcaagga  | ggaggaaaag  | 16380 |
| agaaggctgg  | aaagggacag  | ctcagctggg  | gacacgggag  | tcccctgacc  | tttgtcgggg  | 16440 |
| ggcaggcttg  | ggctcggcga  | tcacaaggaa  | gaggccaagg  | ccgccagtgc  | agaggggaag  | 16500 |
| gaaagagcgc  | ggcagcctta  | gggattttta  | gatgggcagc  | agatgccttt  | aggggtgagag | 16560 |
| atgtacgaag  | agaggacact  | tgtgcccccc  | ccatcatctg  | agaaaaacaa  | cagccagatg  | 16620 |
| tgtccttgcg  | aggtccacct  | tgcccagagc  | tccctcgggg  | actctgtcct  | ggtggcagg   | 16680 |
| ttttgggtacc | ctggccctaga | aggccctctc  | tcactcttcc  | aaggggagg   | gacgcttccg  | 16740 |
| gacggagcct  | tgggtgctccc | tggccgggtg  | tgcctaagg   | ggctctagga  | ggaatcccag  | 16800 |
| agccaagcat  | tactcagagg  | gcgcctggaa  | tgttcccctg  | gaatgctccc  | agcccccca   | 16860 |
| ctggccccaa  | ccactctcac  | aggcccgc    | tgcaggagcc  | aggccccagg  | caccagagc   | 16920 |
| ctgcagcagc  | cctccttccc  | ccggtaccca  | gtcccagctc  | ccagaacaga  | cagcctcccc  | 16980 |
| cctccacgca  | gcccctggcct | cagtccctgt  | gggctgatgg  | ctgcctgtgg  | aagtgactca  | 17040 |
| gctcctgcta  | ggccacccca  | actccttttt  | tctcctccac  | cctctctccc  | agactacaaa  | 17100 |
| catcaaagac  | ccttcctcca  | agaagccctc  | cttgattgga  | tgagtgaatt  | gccatcaggc  | 17160 |
| agatgagggc  | cgagaggagt  | ctgccacctt  | ggaaaggagg  | ctagaggggc  | cagtgcagg   | 17220 |
| agggctctga  | gtggatgtgg  | gggaggggaa  | ggaggggagg  | tctctcagcc  | cagagagcac  | 17280 |
| ttaaactgaga | gtagagaacc  | aagctttgct  | gctcctaggc  | ctctaagggt  | ttggggaaaga | 17340 |
| ggtaggggtg  | gcccgggcac  | aggtgtggtg  | tgggtgcagt  | gtggtgtgtg  | ggtgctgtcc  | 17400 |
| acatggcctt  | gcgcgcacgt  | gctggccacg  | ggcaccctga  | ccccaatgag  | ggagagagg   | 17460 |
| gcagagctgg  | agctggagct  | ggagctccgg  | tgaccgggtg  | aatgggggtg  | gaaccgagg   | 17520 |
| gagccaggct  | ggtattgggc  | acatagacgc  | ccctctccca  | ggggtcccat  | cacctccct   | 17580 |
| gaccccagga  | tagggctcag  | aggggaggga  | gcagtggacc  | gcctggggcc  | ctcccctggg  | 17640 |
| gccagaacag  | accaggcccc  | tgtacctgtt  | tgggtccccac | acagtgtgt   | ggaagccacc  | 17700 |
| gcccagctctg | catagcacag  | cccagccccg  | catgccccct  | ccttggttgc  | cctccctgtt  | 17760 |
| cccggccagg  | cacttgctgt  | gcaggactgg  | ctaactctcc  | caccgcttg   | cagaggttgt  | 17820 |
| tccagcccca  | tcttaacatc  | tttgtttgga  | ggggttacc   | cgaggagaca  | gctgcagtct  | 17880 |
| ttccagagca  | ctgctaaaca  | gacaccttct  | atctggagag  | gcccctctct  | atctcaccaa  | 17940 |
| acaaggcaac  | aataataaca  | acatacacac  | tgccttgctg  | ccctgggagg  | agggacgagg  | 18000 |
| ggtgagcagg  | gtggaggcca  | cagctagtct  | tgcagcctga  | gagcaaagca  | gggactctgg  | 18060 |
| gggactcttg  | ggcatggggg  | cttcctagag  | gatggagccc  | cgctgagtcc  | taaggggtgg  | 18120 |
| aggagcagga  | gcgggtcaca  | cggtggcctg  | cggatggaag  | ctggttgtga  | gagcgagaat  | 18180 |
| ccaggcagag  | ggggctacgg  | ctatgggctg  | ggggctgggg  | gctgggctgt  | ccccgagggg  | 18240 |
| gagggagcca  | tgcctctctg  | tttgccagcg  | gagtggcagc  | cgggcagtgt  | gggcaagtcc  | 18300 |
| gggcccgggg  | ccagcccaag  | cacacttgag  | cgtccctggg  | caggtcccac  | ggagaccccc  | 18360 |
| ccaaagagtc  | cccacgcctt  | gacctactgg  | cogtatgggt  | ccggggccgt  | gagacccctc  | 18420 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gcgcgctgac  | ccgagctctg  | agcagaaccc  | atccccgcc   | ccaccacgc   | gcctagcctg  | 18480 |
| ccccacagg   | cgcacccgc   | ccgcgtcctc  | accttgaagc  | accccggcgc  | ctggcactgg  | 18540 |
| ccagagcagc  | agcagtagta  | gcagcagcag  | caacggggtc  | ccccgagctc  | tccggggcct  | 18600 |
| ccagcccata  | gctgtgagct  | cctcggcctc  | taggcagcgg  | ctcgcaactc  | cggtccgcgc  | 18660 |
| caggctggat  | tgcggccgac  | ccgtgcccg   | tgcagcctca  | ggccgcgcgc  | ttcggaacctt | 18720 |
| cccgccccca  | cctcccaccg  | cccgccctcg  | ctcccgcctc  | ccctcccgc   | caaccccgcct | 18780 |
| cggagcctgg  | ccaggggccc  | cgacggcgcg  | cgccatgggg  | gagccggggtc | gccactcccc  | 18840 |
| gaccgcccgc  | cctcgagggg  | gtggagctgg  | gcggaggagg  | gaatccgtgc  | ggccccctcg  | 18900 |
| atgaccgggc  | cgagccgtcc  | ctccccgtcg  | gtctcagagg  | gcctctactc  | ctgagaggag  | 18960 |
| gagagaaccg  | ctgggaaggt  | tcttgaggga  | cgcggcggtg  | gtgggatgag  | gcgggtgggca | 19020 |
| aaggccgcct  | ctcgctgctg  | aagttaggcc  | caggagcgcg  | atcttccgtg  | gtctcctggg  | 19080 |
| gcgatctct   | gtccccctct  | tgtaccctgt  | cctgccccga  | gggtgccttg  | gcggaggttg  | 19140 |
| agtcgggtca  | tccacctgca  | ctgggtgccc  | ccaaggatag  | gaagggttcag | gcaaccggct  | 19200 |
| gccgctgtct  | tgggggcttc  | attgctgggc  | aaaggcgatg  | cagcagacgg  | agacaacctt  | 19260 |
| tcttccctgg  | cggtagccag  | agggcagaat  | tgcataaaag  | ctgcagactc  | ccaggcctgg  | 19320 |
| gagacccttt  | cggcctcagt  | aacatctgtt  | tcatgtttta  | aacttttgtt  | ttcctactcg  | 19380 |
| gtgcaaat    | ggatgagatg  | ttactttttt  | tttttttttt  | ttttttgaga  | tggagtctcc  | 19440 |
| ctctgtcgcc  | aggtcggagt  | gcagcggcgc  | gatcttggtc  | caactgcaacc | tccgactccc  | 19500 |
| tggttcaagc  | gattctcctg  | cctcagcctc  | ccagtagctg  | gggactacag  | gcgcgcgcta  | 19560 |
| ccacccccag  | ctaacctttt  | tatttttagc  | agagacgagg  | tttcaccatt  | tgggccaagg  | 19620 |
| tgggtctcaat | ctcctgatct  | cgtgatccac  | ccgcctcggc  | ctctcaaagc  | gctggggatta | 19680 |
| caggcatgag  | ccaccgcgcc  | cggccggaga  | tgttaacttt  | taagcaaate  | tttttttttt  | 19740 |
| tttttttttt  | tgagacagag  | tttctctctt  | gttaccacga  | ctggagtgca  | atggcatgat  | 19800 |
| ctgggctcac  | tgaacctct   | gcctcccaga  | ttcaagtgat  | tcttctgctt  | cagcctcccg  | 19860 |
| agtagctggc  | attacaggca  | ttcgccacca  | cgcttggtta  | atthttgtatt | tttagtagag  | 19920 |
| atgggggttt  | tccatgttgg  | tcaggctggg  | ctcgaactcc  | cgacctcagg  | tgatctgccc  | 19980 |
| gcctcggcct  | cccaaagcgc  | tgggaattaca | ggcgtgagac  | accgcaccca  | gcctactttt  | 20040 |
| gcagtaaatct | atthttgttt  | gagaatttgg  | aatgtagtaa  | tttgggttagt | gaaagtctga  | 20100 |
| gcagtggagag | aaacctacat  | tcacatatct  | caaaatcaaa  | aagtacagaa  | agcataggga  | 20160 |
| aaagtctccg  | tgtctcttagc | cctcctcacc  | aacaggaaac  | caatatgatt  | agtttctttc  | 20220 |
| ataggctttt  | agattattht  | ttcacactca  | agacaataca  | gacatattht  | tttctcttat  | 20280 |
| taacgttttt  | ctgcacttht  | atthttcttt  | tttttttggt  | cgcttaatac  | accttagata  | 20340 |
| tcagtgcgtt  | tagagggtcc  | ttgttggtct  | tatgattatt  | atthtagagac | agggtctcac  | 20400 |
| tctgtcaccc  | acgttagagg  | acagtggcct  | gatcatgcct  | cattgcagcc  | ttgaaatcct  | 20460 |
| gggtctcaagg | tatcctccca  | cctcagcctc  | ctgagtagct  | ggaactacag  | gcacacggca  | 20520 |
| ccaggccccag | ctaaaattht  | taatthttct  | gtagacaggg  | ggtctcactt  | tgtttcccg   | 20580 |
| gctggtctca  | aactcctggg  | cttgcccgag  | cgcagtgtct  | catgcctgta  | atcccagcac  | 20640 |
| tttgggagcg  | cagggcgggc  | agatcactgg  | aggtcaggag  | ttcaagacca  | gtctggccaa  | 20700 |
| catggtgaaa  | ccccatctct  | actaaaaata  | caaaaattag  | ccgggcatgg  | tggtagcgcg  | 20760 |
| ctgtagttcc  | agctacttht  | gaggctgagg  | caggaaaatc  | gcttgaaact  | agaagggtgga | 20820 |
| ggttgcagcg  | agccgagatc  | atgccattgc  | actccagcct  | gggcaacaag  | agcgaaactc  | 20880 |
| cgtctcaaaa  | aataaaaaata | aaaataaaaa  | gaactcctga  | tcttaagtga  | tctcctgccc  | 20940 |
| tcagcttctc  | aaatcgctgg  | aattacagga  | gtgagtcacc  | acagctgtcc  | agctacgaga  | 21000 |
| ttattactta  | ttattactac  | tttggttttt  | caaatcaact  | tcattaagggt | ataatthtaca | 21060 |
| cacaataaaa  | tgcacttatt  | tttaagtggc  | agtaagatga  | gtttcgataa  | gtgtatataa  | 21120 |
| ctacataaag  | atcaactataa | tgcagacaca  | ttccctcact  | cacagaaaga  | gcctctgtgc  | 21180 |
| cttcagcca   | aacttcccca  | ctcccaaccc  | cagacagcca  | ctgatctgtt  | gttctctgtc  | 21240 |
| tatagataag  | ttttgcctgt  | tctagaatth  | catataaatg  | gaatcatggc  | gcatgcactc  | 21300 |
| ttctgtgtct  | ggcttccctc  | cctctttccg  | atgtttttga  | gattcattta  | cactatthttg | 21360 |
| catatcaata  | gtttgttctc  | tcgtattgct  | gaatagtgtt  | cggtgggtttg | agggaaccac  | 21420 |
| agtttctcta  | ctcaccagtg  | caccataggg  | ttatthttcca | gttaggggct  | cttataattg  | 21480 |
| gaactatatt  | tgcacagaga  | gagagagagg  | aagaaagagg  | gagagagata  | tttattatag  | 21540 |
| caattggctc  | acgtgattat  | ggaggccaaa  | aagttcccg   | atctgccatc  | tgcaagctgg  | 21600 |
| agaacgagga  | aagccagtg   | tgtgattcag  | tttgagttca  | aaggcctgag  | aaccaggagc  | 21660 |
| accagtatgg  | aggtggctcg  | agctcagaac  | aagttgggga  | caggaaagca  | gagcagcacc  | 21720 |
| ccagagcagc  | cctcagcga   | cacctcttca  | gtaaagcaag  | gctgaacaca  | gaggggctgg  | 21780 |
| cttcagtgtg  | gagtgcagg   | acagaaggca  | gctcagggag  | ctactctggc  | gttcttctgt  | 21840 |
| actggtatct  | ttacctcgaa  | ctggcccaact | cctacttaaa  | ctgcaggcca  | tggctthta   | 21900 |

73/154

```

gtcctgtcat tcagaggctg tcccttaccc aaagccaggt tagcatcccc tgactgacac 21960
ttctccctgc aacacgtttc agaaggccct gtagtcgtcc acttccctgt ctctctcccc 22020
aagctcctga gctccatgtg gtctgggaat atgtgtgttg ctcaactcct agcacagtca 22080
gtgctaataa ctgactgtag aggggacaca gtcgaaaagc cacatgggga tcagagtcac 22140
ccttacacag ttgacacctc ccaaaccag atgagctgtg tccaagtga ggtcagagga 22200
atcttctgcc gaagtctctg agaaagggtt tatttacatt ttgagggtgc aggggaggag 22260
atgaggccat caaaccaaag ctgagggaaga gggatcctag gatgcaccga gcagctccgg 22320
gggcgcctga cagcacctgg gaaagatggc ttctccactg gcttgttggc gtcaccctcc 22380
agaggggcat caggaaatgt cctgggaacc aggcaaacca gtgagcatta acccttagaa 22440
gtgcttggca tgggtgacac ccaccatctg taaacacgac ttctcccaag gagtgaagca 22500
gaacaggatg tctgaggagg gcaactccgac tccagccttc agagatcgcc aggggtggc 22560
ctgggtgacga caggctgatg cttgggtgcc ccagaaaagg tcatgtgtgt gaatgggggc 22620
cccaaagcca acgcttcac cctgacagcc tgggtgcattt agaggggaac tttttgtccc 22680
ttggcaagggt ggggtggaatt tcagggttcat agggcaaggg tatttttagct ttaatagata 22740
ttgtcaaaca gttttccaaa gtcatgtgtac acactctgtg attctactta tgtaaagttt 22800
aaaaacaggc aaatcaaata tatggtgttc gaagtcaaga cagtagttac ccttgtgggg 22860
gctgcaactg gtacagagtg taagggggga ctgtaggatg gtctatttct tgatctgggt 22920
gtgttcgctt ttggaaaagt ccttgagttg catttataat gtgtgaactt ttctgtatgt 22980
tacactttaa ttgaatgtac aaaaagtctc aggaggcctc agaccactgg aagcggacac 23040
aactaaccoc tctgagagcc tccaatccaa tagtgacata tgtccccttg gaagtatgca 23100
gaagcagggt aagactccta agccggatat tcccaaatcc cccagtagc cgcagcttca 23160
gcagctgctt atggtcctcc ctacaccctc tcttccccag acagcccca aacatctggc 23220
tgcatttgac ttgctctctc cctgtcccac ctctggattt agtccatgtt ctccaccctc 23280
cccactgtca gcaatgtaga caagacaaac gcttagttca cgtgccacc tactgcgtgc 23340
catgcacggg gctgggtcatt gtgggtggca aatgtgagca acacacgaag cctcaaggag 23400
cagaaaggga cacaatcac ttcagcgtaa ggtaatttgt gataaatgtc atgtaacttg 23460
cagcccctgg ccccctccta cagatgggtg ctaagaataa accccactaa catgtgactc 23520
ctctgttcta gcccagctgt ttgggttgca agaaagagac tcactccagt tgcg 23574

```

<210> 7  
 <211> 65  
 <212> DNA  
 <213> Homo sapiens

```

<400> 7
agtctccgtg ctcttagccc tcctcacaa caggaaacca atatgattag tttctttcat 60
aggct                                           65

```

<210> 8  
 <211> 656  
 <212> DNA  
 <213> Homo sapiens

```

<400> 8
gtgcagcctc aggcgcgcgc cttcggacct tccgcgcccc acctccacc gccgcacctc 60
gctcccgcct cccctccccg ccaacccgcg tcggagcctg gccaggggcc ccgacggcgc 120
gcgccatggg ggagccgggt cgcactccc ggaccgcgc cctcgaggg ggtggagctg 180
ggcggaggag ggaatccgtg cggcccctcg gatgaccggc ccgagccgtc cctcccctgc 240
ggtctcagag ggcctctact cctgagagga ggagagaacc gctgggaagg ttcttggagg 300
accgcggcgt ggtgggatga ggcggtgggc aaaggccgc tctcgctgct gaagtggcc 360
ccaggagcgc gatcttccgt ggtctcctgg ggccgatctc tgtcccctcc ttgctacccg 420
tcttgcgccg aggggtgcct ggcgagggtt gagtgggtc atccacctgc actgggtgcc 480
ccaaggata ggaaggttca ggcaaccggc tgccgtgtc ttgggggctt cattgctggg 540
caaaggcgat gcagcagac gagacaacct ttcttcctg gcggtggcca gagggcagaa 600
ttgcataaaa gctgcagact ccagggcctg ggagaccctt tcggcctcag taacat 656

```

74/154

<210> 9  
<211> 177  
<212> DNA  
<213> Homo sapiens

<400> 9  
cgggcacggg tgggccgcaa tccagcctgg gccgagccgg agttgcgagc cgctgcctag 60  
aggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120  
tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaag 177

<210> 10  
<211> 80  
<212> DNA  
<213> Homo sapiens

<400> 10  
gacatatccc tgggcagcca gtcaccccgc actgggtcct ggatggacaa ccctggcgca 60  
ccgtcagcct ggaggagccg 80

<210> 11  
<211> 77  
<212> DNA  
<213> Homo sapiens

<400> 11  
gtctcgaagc cagacatggg gctggtggcc ctggaggctg aaggccagga gctcctgctt 60  
gagctggaga agaacca 77

<210> 12  
<211> 79  
<212> DNA  
<213> Homo sapiens

<400> 12  
caggctgctg gccccaggat acatagaaac ccactacggc ccagatgggc agccagtggc 60  
gctggccccc aaccacacg 79

<210> 13  
<211> 119  
<212> DNA  
<213> Homo sapiens

<400> 13  
caggctgctg gccccaggat acatagaaac ccactacggc ccagatgggc agccagtggc 60  
gctggccccc aaccacacg tgagatgctt ccatgggctc tgggatgcac cgccagagg 119

<210> 14  
<211> 77  
<212> DNA  
<213> Homo sapiens

<400> 14

75/154

gatcattgcc actaccaagg gcgagtaagg ggcttccccg actcctgggt agtcctctgc 60  
acctgctctg ggatgag 77

<210> 15  
<211> 190  
<212> DNA  
<213> Homo sapiens

<400> 15  
tggcctgata accctcagca ggaatgccag ctattatctg cgtccctggc caccocgggg 60  
ctccaaggac ttctcaaccc acgagatctt tcggatggag cagctgctca cctggaaagg 120  
aacctgtggc cacagggatc ctgggaacaa agcgggcatg accagccttc ctggtggtcc 180  
ccagagcagg 190

<210> 16  
<211> 66  
<212> DNA  
<213> Homo sapiens

<400> 16  
ggcaggcgag aagcgcgcag gacccggaag tacctggaac tgtacattgt ggcagaccac 60  
accctg 66

<210> 17  
<211> 72  
<212> DNA  
<213> Homo sapiens

<400> 17  
ttcttgactc ggcaccgaaa cttgaaccac accaaacagc gtctcctgga agtcgccaac 60  
tacgtggacc ag 72

<210> 18  
<211> 167  
<212> DNA  
<213> Homo sapiens

<400> 18  
cttctcagga ctctggacat tcaggtggcg ctgaccggcc tggaggtgtg gaccgagcgg 60  
gaccgcagcc gcgtcacgca ggacgccaac gccacgctct gggccttccct gcagtggcgc 120  
cgggggctgt gggcgcagcg gccccacgac tccgcgcagc tgctcac 167

<210> 19  
<211> 85  
<212> DNA  
<213> Homo sapiens

<400> 19  
gggcgcgcgc ttccagggcg ccacagtggg cctggcgccc gtcgagggca tgtgccgcgc 60  
cgagagctcg ggaggcgtga gcacg 85

<210> 20

76/154

<211> 143  
<212> DNA  
<213> Homo sapiens

<400> 20  
gaccactcgg agctcccat cggcgccgca gccaccatgg cccatgagat cggccacagc 60  
ctcggcctca gccacgaccc cgacggctgc tgcgtggagg ctgcggccga gtccggaggc 120  
tgcgtcatgg ctgcggccac cgg 143

<210> 21  
<211> 178  
<212> DNA  
<213> Homo sapiens

<400> 21  
gcacccgttt ccgcgcgtgt tcagcgccctg cagccgcgcg cagctgcgcg ccttcttccg 60  
caagggggggc ggcgcttgcc tctccaatgc cccggacccc ggactcccgg tgccggccggc 120  
gctctgcggg aacggcttcg tggaagcggg cgaggagtgt gactgcggcc ctggccag 178

<210> 22  
<211> 90  
<212> DNA  
<213> Homo sapiens

<400> 22  
gagtgcgcgc acctctgctg ctttgctcac aactgctcgc tgcgcccggg ggcccagtg 60  
gccacggggg actgctgcgt gcgctgcctg 90

<210> 23  
<211> 196  
<212> DNA  
<213> Homo sapiens

<400> 23  
ctgaagccgg ctggagcgt gtgccgccag gccatgggtg actgtgacct ccctgagttt 60  
tgcacggggc cctcctccca ctgtcccca gacgtttacc tactggacgg ctcacctgt 120  
gccaggggca gtggctactg ctgggatggc gcatgtccca cgctggagca gcagtgccag 180  
cagctctggg ggccctg 196

<210> 24  
<211> 107  
<212> DNA  
<213> Homo sapiens

<400> 24  
gtctccaccc agctcccag gcctgtttcc aggtgggtgaa ctctgcggga gatgctcatg 60  
gaaactgcgg ccaggacagc gagggccact tcctgccctg tgcaggg 107

<210> 25  
<211> 199  
<212> DNA  
<213> Homo sapiens

77/154

&lt;400&gt; 25

ggatgccctg tgtgggaagc tgcagtgccca ggggtggaaag cccagcctgc tgcaccgca 60  
catggtgccca gtggactcta ccgttcacct agatggccag gaagtgactt gtcggggagc 120  
cttggcactc cccagtgcgc agctggacct gcttggcctg ggcctggtag agccaggcac 180  
ccagtgtgga cctagaatg 199

&lt;210&gt; 26

&lt;211&gt; 109

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

gtttgcaata gcaaccataa ctgccactgt gctccaggct gggctccacc cttctgtgac 60  
aagccaggct ttggtggcag catggacagt ggcctgtgc aggctgaaa 109

&lt;210&gt; 27

&lt;211&gt; 148

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

accatgacac cttcctgctg gccatgctcc tcagcgtcct gctgcctctg ctcccagggg 60  
ccggcctggc ctggtgttgc taccgactcc caggagccca tctgcagcga tgcagctggg 120  
gctgcagaag ggaccctgcg tgcagtgg 148

&lt;210&gt; 28

&lt;211&gt; 92

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

ccccaaagat ggcccacaca gggaccaccc cctgggcggc gttcacccca tggagttggg 60  
ccccacagcc actggacagc cctggcccct gg 92

&lt;210&gt; 29

&lt;211&gt; 72

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 29

accctgagaa ctctcatgag cccagcagcc accctgagaa gcctctgccca gcagtctcgc 60  
ctgaccccca ag 72

&lt;210&gt; 30

&lt;211&gt; 1031

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 30

cagatcaagt ccagatgccca agatcctgcc tctggtgaga ggtagctcct aaaatgaaca 60  
gatttaaaga caggtggcca ctgacagcca ctccaggaaac ttgaactgca ggggcagagc 120  
cagtgaatca ccggacctcc agcacctgca ggcagcttgg aagtttcttc cccgagtggc 180  
gcttcgaccc acccactcca ggaaccacaga gccacattag aagttcctga gggctggaga 240

78/154

```

acactgctgg gcacactctc cagctcaata aaccatcagt ccagaagca aaggtcacac 300
agccctgac ctccctcacc agtggaggct gggtagtgct ggccatccca aaagggctct 360
gtcctgggag tctgggtgtgt ctctacatg caatttccac ggaccagct ctgtggaggg 420
catgactgct ggccagaagc tagtggctct ggggacctat ggttcgactg agtccacact 480
cccctgcagc ctggctggcc tctgcaaaca aacataattt tggggacctt ccttcctgtt 540
tcttcccacc ctgtcttctc ccctaggtgg ttcctgagcc cccaccccca atcccagtg 600
tacacctgag gttctggagc tcagaatctg acagcctctc cccattctg tgtgtgtcgg 660
ggggacagag ggaaccattt aagaaaagat accaaagtag aagtcaaaag aaagacatgt 720
tggctatagg cgtgggtggct catgcctata atcccagcac tttgggaagc cggggtagga 780
ggatcaccag aggccagcag gtccacacca gcctgggcaa cacagcaaga caccgcatct 840
acagaaaaat tttaaaatta gctgggcgtg tggtgtgtga cctgtaggcc tagctgctca 900
ggaggctgaa gcaggaggat cacttgagcc tgagttcaac actgcagtga gctatgggtg 960
caccactgca ctccagcctg ggtgacagag caagaccctg tctctaaaat aaattttaaa 1020
aagacataaa a 1031

```

<210> 31  
 <211> 78  
 <212> DNA  
 <213> Homo sapiens

```

<400> 31
gtgtgccaga gcaggcgctg caggaagaat gccttcagg agcttcagcg ctgcctgact 60
gcctgccaca gccacggg 78

```

<210> 32  
 <211> 6  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: polyhistidine tag

```

<400> 32
His His His His His His
  1             5

```

<210> 33  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: FLAG epitope tag

```

<400> 33
Asp Tyr Lys Asp Asp Asp Asp Lys
  1             5

```

<210> 34  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence

<220>



79/154

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 34

aactcttgaa atgagaagcg tg

22

&lt;210&gt; 35

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 35

aatatcatgc accatgaccc ac

22

&lt;210&gt; 36

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 36

tggagtaagt attgtaaact at

22

&lt;210&gt; 37

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 37

ggagcttatc ctggattatc ta

22

&lt;210&gt; 38

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 38

agagccacac atccatgtcc tg

22

&lt;210&gt; 39

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 39

aagccactct gtgaattgcc at

22

&lt;210&gt; 40

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 40

gagtagtcgt agtaccagat gg

22

&lt;210&gt; 41

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 41

gtctggcaat ggagcatgaa aa

22

&lt;210&gt; 42

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 42

attagagcac atgaaggaaa gg

22

&lt;210&gt; 43

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 43

acactgcttt gggggacagg ct

22

&lt;210&gt; 44

&lt;211&gt; 22

81/154

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 44  
cacgacgccca cagagccagc tc 22

<210> 45  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 45  
aaccaccacg gattcacgct tc 22

<210> 46  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 46  
ataaccagat ggctgtgggt ca 22

<210> 47  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 47  
atccccgcaa tgaaatagtt ta 22

<210> 48  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 48  
gttgagagcc cacttagata at 22

82/154

<210> 49  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 49  
gcattggggg aagccaggac at 22

<210> 50  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 50  
gccactagga ggcaatggca at 22

<210> 51  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 51  
cgacggcatc acggccatct gg 22

<210> 52  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 52  
tccaggctca ttcattttca tg 22

<210> 53  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 53  
tgacatcaac ttctcctttc ct 22

<210> 54  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 54  
agttgcagag acctagcctg tc 22

<210> 55  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 55  
tctgggagag gacggagctg gc 22

<210> 56  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 56  
tgtaggacta tattgctc 18

<210> 57  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 57  
cgacatttag gtgacact 18

<210> 58  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: BstXI-linker  
adapter

<400> 58  
gttttcacca cgggg

15

<210> 59  
<211> 11  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: BstXI-linker  
adapter

<400> 59  
gtggtgaaga c

11

<210> 60  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic  
peptide

<400> 60  
Asp Pro Gln Ala Asp Gln Val Gln Met  
1 5

<210> 61  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic  
peptide

<400> 61  
Asp Pro Gln Asp Gln Val Gln Met  
1 5

<210> 62  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> (1)..(11)  
<223> "Xaa" represents a variable amino acid

<220>  
<223> Description of Artificial Sequence: Zn-binding

85/154

consensus sequence

&lt;400&gt; 62

His Glu Xaa Xaa His Xaa Xaa Gly Xaa Xaa His  
1 5 10

&lt;210&gt; 63

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 63

ctgcctagag gccgagga

18

&lt;210&gt; 64

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 64

caggagacca cggaagatcg

20

&lt;210&gt; 65

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 65

ttgcctgaac cttcctatcc

20

&lt;210&gt; 66

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 66

cccctgtggt cctcaggtc

19

&lt;210&gt; 67

&lt;211&gt; 20

&lt;212&gt; DNA

86/154

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 67

gctccacact ctttcttgcc

20

&lt;210&gt; 68

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 68

aggcaggagg aagctgaat

19

&lt;210&gt; 69

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 69

cctaccacac cctccctctt

20

&lt;210&gt; 70

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 70

cctacccctc tgcacccta

19

&lt;210&gt; 71

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 71

aacttccttc tgggagctgg

20

&lt;210&gt; 72



87/154

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 72  
cacaccctgg tgaggagaga 20

<210> 73  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 73  
ccacgaagga ccaccg 16

<210> 74  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 74  
ctcacgtggg tgccctctg 18

<210> 75  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 75  
ctctacggcc gcagtgac 18

<210> 76  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 76  
gtccctccat gcccaatg 18

<210> 77  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 77  
cagggttaagt cggctcgc 18

<210> 78  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 78  
ctctctctgc cttccccac 19

<210> 79  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 79  
tctactgtgg ggaagatggg 20

<210> 80  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 80  
cccctctact tcctcccca 19

<210> 81  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 81

89/154

gaccttgggg ttccataatcc

20

&lt;210&gt; 82

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 82

gtgcacctgc tcaggactc

19

&lt;210&gt; 83

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 83

cctggactct tatcacgttg c

21

&lt;210&gt; 84

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 84

ttaccctcca ccatttctcc

20

&lt;210&gt; 85

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 85

gtggagaggg aagggaag

20

&lt;210&gt; 86

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

<400> 86  
cccatgggt tgaatttaca 20

<210> 87  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 87  
gcagctaggc ctacaggtac a 21

<210> 88  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 88  
accacgccta tagccaacat 20

<210> 89  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 89  
aggtgtagca ctgggattgg 20

<210> 90  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 90  
ccccaggacc actagcttct 20

<210> 91  
<211> 20  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 91

attgagctgg agagtgtgcc

20

&lt;210&gt; 92

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 92

ttcaagttcc tggagtggct

20

&lt;210&gt; 93

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 93

acaaggaccc totaaacgca

20

&lt;210&gt; 94

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 94

acccttctgt gacaagccag

20

&lt;210&gt; 95

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 95

gtgttgctac cgactcccag

20

&lt;210&gt; 96

&lt;211&gt; 18

&lt;212&gt; DNA

92/154

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 96

cccaggtgca gagagcag

18

&lt;210&gt; 97

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 97

gctcctcttg tccactctcc t

21

&lt;210&gt; 98

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 98

gccacttcct ctgcacaaat

20

&lt;210&gt; 99

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 99

ttctctgtga cctgggtggt

20

&lt;210&gt; 100

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 100

atttgggccca gagatggg

18

&lt;210&gt; 101

<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 101  
ggcagaggag caaggtgg 18

<210> 102  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 102  
atggcttgga atcatcaagg 20

<210> 103  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 103  
tagagagagg aggtgccagc 20

<210> 104  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 104  
aaagatggcc cacacagg 18

<210> 105  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 105  
agaactctca tgagcccagc 20

<210> 106  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 106  
agctctgagc agaacccatc 20

<210> 107  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 107  
ctcgaggggg tggagctg 18

<210> 108  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 108  
gagaggagga gagaaccgct 20

<210> 109  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 109  
agtgacttgg tggttctggg 20

<210> 110  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 110



95/154

tgtcatctgc accctctctg

20

&lt;210&gt; 111

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 111

aagagggagg gtgtggtagg

20

&lt;210&gt; 112

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 112

gtgatcaggc cactaggggtg

20

&lt;210&gt; 113

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 113

atacagcatt ccactccca

20

&lt;210&gt; 114

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 114

gaaggcagaa atcccggt

18

&lt;210&gt; 115

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

<400> 115  
caccagcacc tgccctgtc 18

<210> 116  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 116  
gggtcagagg cacccac 17

<210> 117  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 117  
gccgtagagc ctcctgtct 19

<210> 118  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 118  
gacgaccaaa gaaacgcag 19

<210> 119  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 119  
tgagcggaga gggcaagt 18

<210> 120  
<211> 20  
<212> DNA  
<213> Artificial Sequence

97/154

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 120

aaaccctcac cctgaacctt

20

&lt;210&gt; 121

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 121

aagggtgctc gtgtcctct

19

&lt;210&gt; 122

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 122

ccactcagct ccactcccta

20

&lt;210&gt; 123

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 123

ggattcaaac ggcaaggag

19

&lt;210&gt; 124

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 124

gctgagtcct gagcaggtg

19

&lt;210&gt; 125

&lt;211&gt; 19

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 125

gaaccgcagg agtaggctc

19

<210> 126

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 126

atatggtcag caggagaccc

20

<210> 127

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 127

gcctcctggg ctccatgata a

21

<210> 128

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 128

gaggctttga atccagggtcc

20

<210> 129

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 129

cagcaagaca ccgcatctac

20

<210> 130

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 130  
gggacagagg gaaccattta 20

<210> 131  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 131  
ttccttctctg tttcttccca 20

<210> 132  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 132  
gtcctgggag tctggtgtgt 20

<210> 133  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 133  
aggaacccag agccacacta 20

<210> 134  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 134  
tgcctctggt gagaggtagc 20

<210> 135  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 135  
ttcctggatc actggtcctc 20

<210> 136  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 136  
ttcgagcagt gagagaaacc t 21

<210> 137  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 137  
ctgggagtcg gtagcaaca 19

<210> 138  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 138  
aggccactgg aacctcct 18

<210> 139  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 139

101/154

gcagcatggt acagggactg

20

&lt;210&gt; 140

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 140

cagctgacca gtggtatgga

20

&lt;210&gt; 141

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 141

tgtcagacat ggccacagag

20

&lt;210&gt; 142

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 142

agggtcctct tagctgccac

20

&lt;210&gt; 143

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 143

aggccttgtc atttctgtg

20

&lt;210&gt; 144

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

<400> 144  
caaagaacct tggatgtccg 20

<210> 145  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 145  
ctcagctccc ttcttgctc 19

<210> 146  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 146  
ctgtgtgggc catctttg 18

<210> 147  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 147  
ggagaaatgg tggagggtaa 20

<210> 148  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 148  
aaagccacag cttctccct 19

<210> 149  
<211> 20  
<212> DNA  
<213> Artificial Sequence



&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 149

aggtttcttg gctcaggtta

20

&lt;210&gt; 150

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 150

gtaggtgtgc cagagcagg

19

&lt;210&gt; 151

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 151

tgtggaccta gaatggtgag c

21

&lt;210&gt; 152

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 152

caaagtcaca caacaagcgg

20

&lt;210&gt; 153

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 153

caggatcttg gcattctggac

20

&lt;210&gt; 154

&lt;211&gt; 20

&lt;212&gt; DNA

104/154

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 154

ctggcttgtc acagaagggt

20

&lt;210&gt; 155

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 155

ctggagcaca gtggcagtta

20

&lt;210&gt; 156

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 156

tttggtcgtc cctcagtttc

20

&lt;210&gt; 157

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 157

cctctcagga gtagaggccc

20

&lt;210&gt; 158

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 158

agcggttctc tcctcctctc

20

&lt;210&gt; 159

105/154

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 159  
cctctcagga gtagaggccc 20

<210> 160  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 160  
atgttactga ggccgaaagg 20

<210> 161  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 161  
ccctttccag ccttctcttt 20

<210> 162  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 162  
caggactgca aacatcctga 20

<210> 163  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 163  
tccctgggtgc ttcccata 18

<210> 164  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 164  
aggcaggagg aagctgaat 19

<210> 165  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 165  
cctcttgccc ctcttgct 18

<210> 166  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 166  
cctgaatgtc cagagtcctg a 21

<210> 167  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 167  
ggcctcgagt ccagtatatt 20

<210> 168  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 168

107/154

agagcctcct gtctctccct

20

&lt;210&gt; 169

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 169

tcgccctcag cttctcag

18

&lt;210&gt; 170

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 170

tcacgtgggt gcctctga

18

&lt;210&gt; 171

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 171

gggttacttc ccctctctgg

20

&lt;210&gt; 172

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 172

ctgggctttc caccctgg

18

&lt;210&gt; 173

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

<400> 173  
ctgggctttc caccctgg 18

<210> 174  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 174  
tccaggtggt gaactctgc 19

<210> 175  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 175  
tagaatggtg agctctgccc 20

<210> 176  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 176  
gaccttgagg ttccataatcc 20

<210> 177  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 177  
ccaagcacac ttgagcgtc 19

<210> 178  
<211> 18  
<212> DNA  
<213> Artificial Sequence

109/154

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 178

agccatgccc tctgcttt

18

&lt;210&gt; 179

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 179

cagcccaagc acattga

18

&lt;210&gt; 180

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 180

cccatagctg tgagctcctc

20

&lt;210&gt; 181

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 181

aaagcttcag gaccacaaa

20

&lt;210&gt; 182

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 182

atcttggtcc ctgccattc

19

&lt;210&gt; 183

&lt;211&gt; 18

&lt;212&gt; DNA

110/154

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 183

gagggagctc tttcccca

18

&lt;210&gt; 184

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 184

ggaccaccag gaaggctg

18

&lt;210&gt; 185

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 185

aaccacagct ccagaag

18

&lt;210&gt; 186

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 186

ctgtcacct ggaaaggaac

20

&lt;210&gt; 187

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 187

actgcaggaa ggcccagag

19

&lt;210&gt; 188



111/154

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 188  
accgaaactt gaaccacacc 20

<210> 189  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 189  
tgagggacga ccaaagaaac 20

<210> 190  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 190  
caaagtcaca caacaagcgg 20

<210> 191  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 191  
gaacctgagg gcaccaatta 20

<210> 192  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 192  
ttggccttag ttaattggtg c 21

<210> 193  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 193  
ttggccttag ttaattggtg c 21

<210> 194  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 194  
ctggagcaca gtggcagtta 20

<210> 195  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 195  
aggagtaggc tcaggaagca 20

<210> 196  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 196  
tgtactggga ggtagagggc 20

<210> 197  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 197

agaggggtgac ttggagcaga

20

<210> 198

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 198

aggcaataac ccactcagga

20

<210> 199

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 199

cccatggggtt gaatttacat a

21

<210> 200

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 200

gcctctggtg atcctcctac

20

<210> 201

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 201

actcagtcga accatagggc

20

<210> 202

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 202  
tgtgtgacct ttgcttctgg 20

<210> 203  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 203  
gcatgaagca atgggagaat 20

<210> 204  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 204  
actcagtcga accatagggc 20

<210> 205  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 205  
gcaggaaggt gtcatggtct 20

<210> 206  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 206  
gcaggaaggt gtcatggtct 20

<210> 207  
<211> 18  
<212> DNA  
<213> Artificial Sequence

115/154

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 207

gggcattgga gaggcaag

18

&lt;210&gt; 208

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 208

tctgcctccc agattcaagt

20

&lt;210&gt; 209

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 209

agaatgcctt ccaggagctt

20

&lt;210&gt; 210

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 210

gtgttgctac cgactcccag

20

&lt;210&gt; 211

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 211

ctgcttcttg agcctactcc

20

&lt;210&gt; 212

&lt;211&gt; 19

&lt;212&gt; DNA

116/154

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 212

aacaggaggt tccagtggc

19

&lt;210&gt; 213

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 213

agcgagttgt gattgagggc

20

&lt;210&gt; 214

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 214

tgtgcaggct gaaagtatgc

20

&lt;210&gt; 215

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 215

gccacttcct ctgcacaaat

20

&lt;210&gt; 216

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 216

ctgagcccag aaacctgatt

20

&lt;210&gt; 217

<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 217  
gtgagtgagg caccaggg 18

<210> 218  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 218  
cctagatggc caggaagtga 20

<210> 219  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 219  
ccagaaacct gattaggggg 20

<210> 220  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 220  
tacctctcac cagaggcagg 20

<210> 221  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 221  
gccagaagct agtggtcctg 20

<210> 222  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 222  
gcaggcagct tggaagttt 19

<210> 223  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 223  
ttatcatgga gaccaggatg c 21

<210> 224  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 224  
gacctggatt caaagcctcc 20

<210> 225  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 225  
atgttggtcta taggcgtggt 20

<210> 226  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 226



119/154

ttatcatgga gaccaggatg c

21

&lt;210&gt; 227

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 227

ctgagtggag ggagcagaag

20

&lt;210&gt; 228

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 228

ctgagtggag ggagcagaag

20

&lt;210&gt; 229

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 229

ccatgagatc ggccacag

18

&lt;210&gt; 230

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 230

atttcaaggc tgcaatgagg

20

&lt;210&gt; 231

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

<400> 231  
acttctttcc atggcctctg 20

<210> 232  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 232  
accacccagg tcacagagaa 20

<210> 233  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 233  
tcccaagacc aggctatgtc 20

<210> 234  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 234  
ctggggatga gaagcagc 18

<210> 235  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 235  
cttctccctt ccctctccac 20

<210> 236  
<211> 20  
<212> DNA  
<213> Artificial Sequence

121/154

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 236

atttgtgcag aggaagtggc

20

&lt;210&gt; 237

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 237

catttctctcc aggctctgac

20

&lt;210&gt; 238

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 238

tcagagcctg gaggaaatgt

20

&lt;210&gt; 239

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 239

gttctctggag tgggtgggt

19

&lt;210&gt; 240

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 240

ctgggagtcg gtagcaaca

19

&lt;210&gt; 241

&lt;211&gt; 41

&lt;212&gt; DNA

122/154

&lt;213&gt; Homo sapiens

&lt;400&gt; 241

gccctctgag accgacgggg agggacggct cgggccggtc a

41

&lt;210&gt; 242

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 242

caagaacctt ccagcgggtt ctctctctct ctcaggagta g

41

&lt;210&gt; 243

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 243

caccatctca gctccacact cttctctgcc caggtctcga a

41

&lt;210&gt; 244

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 244

ccaccatctc agctccacac tctttcttgc ccaggtctcg a

41

&lt;210&gt; 245

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 245

acaactaagc catcaccaag gctccttct ctagcccca g

41

&lt;210&gt; 246

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 246

tggtgettcc catattcaca tctcccacaa ctaagccatc a

41

&lt;210&gt; 247

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 247

caggatacat agaaacccac tacggcccag atgggcagcc a

41

<210> 248  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 248  
ccctccaaat cagaagagac aggaattcac aggctcgag t 41

<210> 249  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 249  
agctgctcac ctggaaagga acctgtggcc acagggatcc t 41

<210> 250  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 250  
acttccttct gggagctggg gttgggggtc agggctcaag c 41

<210> 251  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 251  
ttcctgcagt ggcgccgggg gctgtgggag cagcggcccc a 41

<210> 252  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 252  
ggttcagggg gaggggtttcg gggagcttgg gagccggcct g 41

<210> 253  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 253  
cagagaagcg cggggggttg gggactgtcc ctccatgcc a 41

<210> 254  
<211> 41

124/154

<212> DNA  
<213> Homo sapiens

<400> 254  
cccctctctg ggctctgcgc gtctggcggc tgtagccaag c 41

<210> 255  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 255  
cagccgccgc cagctgcgcg ccttcttccg caaggggggc g 41

<210> 256  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 256  
agtggcctcc cagtcaagcg aggggggtgga tccctgcccc a 41

<210> 257  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 257  
tgctggccat gctcctcagc gtctgtctgc ctctgtctcc a 41

<210> 258  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 258  
ctgtcgcttc tgetcccagg ggccggcctg gcctgggtgtt g 41

<210> 259  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 259  
gaagtagctt tgaacaggag gttccagtgg cctcccagtc a 41

<210> 260  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 260

125/154

gcctctgtct caccagtttt cggccctttg ccacttcctc t 41

<210> 261  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 261  
acaaatcacc tctgtcaccc ccttgaagtt cccaaatgct g 41

<210> 262  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 262  
tccataccac tggtcagctg cgggtgctggc tgccctgtg c 41

<210> 263  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 263  
ggtgctggct gcccctgtgc cagggccctg ccttaaccca g 41

<210> 264  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 264  
ggaaatgaca aggccttggg ggatgggatg gggacagtca a 41

<210> 265  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 265  
agggtcatg cctcctgcct ccttccagat gggcagcacc c 41

<210> 266  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 266  
gcccctcccc agccccaggg tctcctgctg accatattca c 41

<210> 267

126/154

<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 267  
cctgggcggc gttcacccca tggagttggg cccacagcc a 41

<210> 268  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 268  
gccccacagc cactggacag ccctggcccc tgggtgagtg a 41

<210> 269  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 269  
gccctggccc ctgggtgagt gaggcaccag ggggaggtgg a 41

<210> 270  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 270  
tgcagcctgg ggccccagtc cttaggggac aacatatacct c 41

<210> 271  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 271  
cactgagtga ggatgggctc tctgccacac agcttgagc c 41

<210> 272  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 272  
ctggtcctca ctgagtgagg atgggctctc tgccacacag c 41

<210> 273  
<211> 41  
<212> DNA  
<213> Homo sapiens



127/154

<400> 273  
atgacctctt ggttatcatg gagaccagga tgctggaagc c 41

<210> 274  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 274  
agcaagacac cgcctctaca gaaaaatttt aaaattagct g 41

<210> 275  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 275  
ggaggatcac cagaggccag caggccaca ccagcctggg c 41

<210> 276  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 276  
atcccagcac tttgggaagc cggggtagga ggatcaccag a 41

<210> 277  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 277  
agcctggctg gcctctgcaa acaaacataa ttttggggac c 41

<210> 278  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 278  
actgagtcca cactccctg cagcctggct ggctctgca a 41

<210> 279  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 279  
tccaggaacc cagagccaca ttagaagttc ctgagggctg g 41

128/154

<210> 280  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 280  
tttttccccg agtggagctt cgacccaccc actccaggaa c 41

<210> 281  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 281  
tcctcattct cagcagatca agtccagatg ccaagatcct g 41

<210> 282  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 282  
ctgaggacca cacgggggtgg tggttggcgg ggtgggtggtt g 41

<210> 283  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 283  
ggctggcagg cagagcctag atggcagcca gagccccagg c 41

<210> 284  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 284  
ctttgctctg tcactcctgc ctcccttggg cgttcacatt c 41

<210> 285  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 285  
gtgagctctg cccacccgac ccctccttgc cgtttgaatc c 41

<210> 286  
<211> 41  
<212> DNA  
<213> Homo sapiens

129/154

<400> 286  
tggcgagggtt actcctacac cgggaggagc accgtcgggt c 41

<210> 287  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 287  
ggctgctcac tattggggcc gcatcgtccc ctgtcccgt t 41

<210> 288  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 288  
gccgcatcgt cccctgtccc gcttggtgtg tgactttgag c 41

<210> 289  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 289  
gccgtccac cccgtcg 17

<210> 290  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 290  
cctcctctct tggcgac 17

<210> 291  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 291  
tccacactct ttcttgcc 18

130/154

<210> 292  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 292  
gctccacact ctttcttgcc 20

<210> 293  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 293  
tcaccaaggc tccttcct 18

<210> 294  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 294  
cagaagagac aggaattcac a 21

<210> 295  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 295  
tggaaggaa cctgtggcc 19

<210> 296  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 296

gggtttcggg gagcttg

17

<210> 297

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 297

gggttggggg actgtc

16

<210> 298

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 298

ctctgcgcgt ctggcg

16

<210> 299

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 299

gccgtccctc cccgtcg

17

<210> 300

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 300

tctctctcta ttggcgaccc

20

<210> 301

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 301  
ctccacactt tttcttgccc a 21

<210> 302  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 302  
gctccacact ctttcttg 19

<210> 303  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 303  
tcaccaagcc tccttcct 18

<210> 304  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 304  
agaagagacg ggaattcac 19

<210> 305  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 305  
tggaaggag cctgtgg 17

<210> 306  
<211> 19  
<212> DNA  
<213> Artificial Sequence

133/154

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 306

agggtttcgt ggagcttg

19

&lt;210&gt; 307

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 307

ggggttggag gactgtcc

18

&lt;210&gt; 308

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 308

gctctgcgca tctggcgg

18

&lt;210&gt; 309

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 309

agtcaagcga ggggggtg

18

&lt;210&gt; 310

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 310

cctcagcgtc ctgctg

16

&lt;210&gt; 311

&lt;211&gt; 18

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 311

aacaggaggt tccagtgg

18

<210> 312

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 312

accagttttc ggcccttt

18

<210> 313

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 313

ctgtcacccc cttgaagt

18

<210> 314

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 314

tcagctgcgg tgctgg

16

<210> 315

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 315

gccttggggg atgga

15

<210> 316



135/154

<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 316  
tcctgcctcc ttccag 16

<210> 317  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 317  
actggacagc cctggc 16

<210> 318  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 318  
ctgtgtggca gagagccca 19

<210> 319  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 319  
aattatgttt gtttgcagag gc 22

<210> 320  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 320  
gaacttctag tgtggctct 19

<210> 321  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 321  
ccaagggagg caggagt 17

<210> 322  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 322  
agtcaagcgt gggggtgg 18

<210> 323  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 323  
ctcctcagca tctgctgc 19

<210> 324  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 324  
gaacaggagt ttccagtggc 20

<210> 325  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 325

137/154

caccagtttt tggccctttg 20

<210> 326  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 326  
ctgtcaccca cttgaagttc 20

<210> 327  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 327  
ggtcagctgt ggtgctgg 18

<210> 328  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 328  
aggccttggg agatgggat 19

<210> 329  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 329  
tcctgcttc ttccag 16

<210> 330  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 330  
actggacagt cctggc 16

<210> 331  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 331  
tgtggcaggg agccca 16

<210> 332  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 332  
attatgtttg cttgcagagg 20

<210> 333  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 333  
ggaacttcta atgtggctct g 21

<210> 334  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 334  
cccaagggaa gcaggagtga 20

<210> 335  
<211> 55  
<212> PRT  
<213> Homo sapiens

139/154

&lt;400&gt; 335

Cys Cys Phe Ala His Asn Cys Ser Leu Arg Pro Gly Ala Gln Cys Ala  
 1 5 10 15

His Gly Asp Cys Cys Val Arg Cys Leu Leu Lys Pro Ala Gly Ala Leu  
 20 25 30

Cys Arg Gln Ala Met Gly Asp Cys Asp Leu Pro Glu Phe Cys Thr Gly  
 35 40 45

Thr Ser Ser His Cys Pro Pro  
 50 55

&lt;210&gt; 336

&lt;211&gt; 11

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 336

Thr Met Ala His Glu Ile Gly His Ser Leu Gly  
 1 5 10

&lt;210&gt; 337

&lt;211&gt; 86

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 337

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
 1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
 20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
 35 40 45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp  
 50 55 60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu  
 65 70 75 80

Leu Glu Lys Asn His Arg  
 85

&lt;210&gt; 338

&lt;211&gt; 48

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 338

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
 1 5 10 15

140/154

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
35 40 45

&lt;210&gt; 339

&lt;211&gt; 178

<212> PRT

<213> Homo sapiens

<400> 339

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
35 40 45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp  
50 55 60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu  
65 70 75 80

Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His  
85 90 95

Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Val  
100 105 110

Arg Cys Phe His Gly Leu Trp Asp Ala Pro Pro Glu Asp His Cys His  
115 120 125

Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys  
130 135 140

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Cys | Ser | Gly | Met | Ser | Gly | Leu | Ile | Thr | Leu | Ser | Arg | Asn | Ala | Ser |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |

Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr  
165 170 175

His Glu

<210> 340

<211> 113

&lt;212&gt; PRT

<213> Homo sapiens

<400> 340

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
1 5 10 15

141/154

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
                     20                    25                    30  
 Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
                     35                    40                    45  
 Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp  
                     50                    55                    60  
 Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu  
                     65                    70                    75                    80  
 Leu Glu Lys Asn His Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser Tyr  
                     85                    90                    95  
 Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr His  
                     100                    105                    110  
 Glu

<210> 341  
 <211> 165  
 <212> PRT  
 <213> Homo sapiens

<400> 341  
 Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
   1                    5                    10                    15  
 Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
                     20                    25                    30  
 Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
                     35                    40                    45  
 Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp  
                     50                    55                    60  
 Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu  
                     65                    70                    75                    80  
 Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His  
                     85                    90                    95  
 Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Asp  
                     100                    105                    110  
 His Cys His Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val  
                     115                    120                    125  
 Val Leu Cys Thr Cys Ser Gly Met Ser Gly Leu Ile Thr Leu Ser Arg  
                     130                    135                    140  
 Asn Ala Ser Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp  
                     145                    150                    155                    160

142/154

Phe Ser Thr His Glu  
165

<210> 342  
<211> 168  
<212> PRT  
<213> Homo sapiens

<400> 342  
Leu Ala Pro Gly Tyr Ile Glu Thr His Tyr Gly Pro Asp Gly Gln Pro  
1 5 10 15  
Val Val Leu Ala Pro Asn His Thr Asp His Cys His Tyr Gln Gly Arg  
20 25 30  
Val Arg Gly Phe Pro Asp Ser Trp Val Val Leu Cys Thr Cys Ser Gly  
35 40 45  
Met Ser Gly Leu Ile Thr Leu Ser Arg Asn Ala Ser Tyr Tyr Leu Arg  
50 55 60  
Pro Trp Pro Pro Arg Gly Ser Lys Asp Phe Ser Thr His Glu Ile Phe  
65 70 75 80  
Arg Met Glu Gln Leu Leu Thr Trp Lys Gly Thr Cys Gly His Arg Asp  
85 90 95  
Pro Gly Asn Lys Ala Gly Met Thr Ser Leu Pro Gly Gly Pro Gln Ser  
100 105 110  
Arg Gly Arg Arg Lys Ala Arg Arg Thr Arg Lys Tyr Leu Glu Leu Tyr  
115 120 125  
Ile Val Ala Asp His Thr Leu Phe Leu Thr Arg His Arg Asn Leu Asn  
130 135 140  
His Thr Lys Gln Arg Leu Leu Glu Val Ala Asn Tyr Val Asp Gln Leu  
145 150 155 160  
Leu Arg Thr Leu Asp Ile Gln Val  
165

<210> 343  
<211> 167  
<212> PRT  
<213> Homo sapiens

<400> 343  
Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
1 5 10 15  
Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe  
20 25 30  
Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp  
35 40 45



143/154

Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly  
 50 55 60

Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met  
 65 70 75 80

Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys  
 85 90 95

Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu  
 100 105 110

Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Asn  
 115 120 125

Ser Asn His Asn Cys His Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys  
 130 135 140

Asp Lys Pro Gly Phe Gly Gly Ser Met Asp Ser Gly Pro Val Gln Ala  
 145 150 155 160

Glu Asn His Asp Thr Phe Leu  
 165

<210> 344  
 <211> 193  
 <212> PRT  
 <213> Homo sapiens

<400> 344  
 Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
 1 5 10 15

Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe  
 20 25 30

Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp  
 35 40 45

Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly  
 50 55 60

Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met  
 65 70 75 80

Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys  
 85 90 95

Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu  
 100 105 110

Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln  
 115 120 125

Ser Arg Arg Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu  
 130 135 140

144/154

Thr Ala Cys His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His  
 145 150 155 160  
 Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly  
 165 170 175  
 Gly Ser Met Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe  
 180 185 190

Leu

<210> 345  
 <211> 126  
 <212> PRT  
 <213> Homo sapiens

<400> 345  
 Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
 1 5 10 15  
 Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala  
 20 25 30  
 Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val  
 35 40 45  
 Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg  
 50 55 60  
 Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys  
 65 70 75 80  
 His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro  
 85 90 95  
 Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met  
 100 105 110  
 Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu  
 115 120 125

<210> 346  
 <211> 93  
 <212> PRT  
 <213> Homo sapiens

<400> 346  
 Ala Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser  
 1 5 10 15  
 Trp Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro  
 20 25 30  
 His Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro  
 35 40 45

145/154

Thr Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu  
 50 55 60

Pro Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro  
 65 70 75 80

Gln Ala Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp  
 85 90

<210> 347

<211> 236

<212> PRT

<213> Homo sapiens

<400> 347

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
 1 5 10 15

Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala  
 20 25 30

Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val  
 35 40 45

Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg  
 50 55 60

Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys  
 65 70 75 80

His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro  
 85 90 95

Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met  
 100 105 110

Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala  
 115 120 125

Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala  
 130 135 140

Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp  
 145 150 155 160

Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His  
 165 170 175

Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr  
 180 185 190

Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro  
 195 200 205

Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln  
 210 215 220

146/154

Ala Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp  
 225 230 235

&lt;210&gt; 348

&lt;211&gt; 302

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 348

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
 1 5 10 15

Gln Gln Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe  
 20 25 30

Gln Val Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp  
 35 40 45

Ser Glu Gly His Phe Leu Pro Cys Ala Gly Arg Asp Ala Leu Cys Gly  
 50 55 60

Lys Leu Gln Cys Gln Gly Gly Lys Pro Ser Leu Leu Ala Pro His Met  
 65 70 75 80

Val Pro Val Asp Ser Thr Val His Leu Asp Gly Gln Glu Val Thr Cys  
 85 90 95

Arg Gly Ala Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu  
 100 105 110

Gly Leu Val Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln  
 115 120 125

Ser Arg Arg Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu  
 130 135 140

Thr Ala Cys His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His  
 145 150 155 160

Cys Ala Pro Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly  
 165 170 175

Gly Ser Met Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe  
 180 185 190

Leu Leu Ala Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala  
 195 200 205

Gly Leu Ala Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg  
 210 215 220

Cys Ser Trp Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp  
 225 230 235 240

Gly Pro His Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu  
 245 250 255

147/154

Gly Pro Thr Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser  
 260 265 270

His Glu Pro Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro  
 275 280 285

Asp Pro Gln Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp  
 290 295 300

<210> 349

<211> 235

<212> PRT

<213> Homo sapiens

<400> 349

Ser Gly Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys  
 1 5 10 15

Gln Gln Leu Trp Gly Pro Asp Gly Gln Glu Val Thr Cys Arg Gly Ala  
 20 25 30

Leu Ala Leu Pro Ser Ala Gln Leu Asp Leu Leu Gly Leu Gly Leu Val  
 35 40 45

Glu Pro Gly Thr Gln Cys Gly Pro Arg Met Val Cys Gln Ser Arg Arg  
 50 55 60

Cys Arg Lys Asn Ala Phe Gln Glu Leu Gln Arg Cys Leu Thr Ala Cys  
 65 70 75 80

His Ser His Gly Val Cys Asn Ser Asn His Asn Cys His Cys Ala Pro  
 85 90 95

Gly Trp Ala Pro Pro Phe Cys Asp Lys Pro Gly Phe Gly Gly Ser Met  
 100 105 110

Asp Ser Gly Pro Val Gln Ala Glu Asn His Asp Thr Phe Leu Leu Ala  
 115 120 125

Met Leu Leu Ser Val Leu Leu Pro Leu Leu Pro Gly Ala Gly Leu Ala  
 130 135 140

Trp Cys Cys Tyr Arg Leu Pro Gly Ala His Leu Gln Arg Cys Ser Trp  
 145 150 155 160

Gly Cys Arg Arg Asp Pro Ala Cys Ser Gly Pro Lys Asp Gly Pro His  
 165 170 175

Arg Asp His Pro Leu Gly Gly Val His Pro Met Glu Leu Gly Pro Thr  
 180 185 190

Ala Thr Gly Gln Pro Trp Pro Leu Asp Pro Glu Asn Ser His Glu Pro  
 195 200 205

Ser Ser His Pro Glu Lys Pro Leu Pro Ala Val Ser Pro Asp Pro Gln  
 210 215 220

Asp Gln Val Gln Met Pro Arg Ser Cys Leu Trp  
 225 230 235

<210> 350  
 <211> 339  
 <212> DNA  
 <213> Homo sapiens

<400> 350  
 cgggcacggg tcggccgcaa tccagcctgg gcgagaccgg agttgcgagc cgctgcctag 60  
 aggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120  
 tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaaggac 180  
 atatccctgg gcagccagtc accccgcact gggtcctgga tggacaaccc tggcgccaccg 240  
 tcagcctgga ggagccggtc tcgaagccag acatggggct ggtggccctg gaggctgaag 300  
 gccaggagct cctgcttgag ctggagaaga accacaggc 339

<210> 351  
 <211> 225  
 <212> DNA  
 <213> Homo sapiens

<400> 351  
 cgggcacggg tcggccgcaa tccagcctgg gcgagaccgg agttgcgagc cgctgcctag 60  
 aggccgagga gctcacagct atgggctgga ggccccggag agctcggggg accccgttgc 120  
 tgctgctgct actactgctg ctgctctggc cagtgccagg cgccgggggtg cttcaaggac 180  
 atatccctgg gcagccagtc accccgcact gggtcctgga tggac 225

<210> 352  
 <211> 562  
 <212> DNA  
 <213> Homo sapiens

<400> 352  
 gcctagaggc cgaggagctc acagctatgg gctggaggcc ccggagagct cgggggaccc 60  
 cgttgctgct gctgctacta ctgctgctgc tctggccagt gccaggcgcc ggggtgcttc 120  
 aaggacatat ccctgggcag ccagtcaccc cgcactgggt cctggatgga caaccctggc 180  
 gcaccgtcag cctggaggag ccggtctcga agccagacat ggggctggtg gccctggagg 240  
 ctgaaggcca ggagctcctg cttgagctgg agaagaacca caggctgctg gcccaggat 300  
 acatagaaac ccactacggc ccagatgggc agccagtggg gctggccccc aaccacacgg 360  
 tgagatgctt ccattgggctc tgggatgcac cgccagagga tcattgccac taccaagggc 420  
 gagtaagggg cttccccgac tcctgggtag tcctctgcac ctgctctggg atgagtggcc 480  
 tgatcaccct cagcaggaat gccagctatt atctgcgtcc ctggccaccc cggggctcca 540  
 aggacttctc aaccacgag at 562

<210> 353  
 <211> 362  
 <212> DNA  
 <213> Homo sapiens

<400> 353  
 gaggccgagg agctcacagc tatgggctgg agggccccga gagctcgggg gaccccggtg 60  
 ctgctgctgc tactactgct gctgctctgg ccagtgccag gcgccggggg gcttcaagga 120  
 catatccctg ggcagccagt caccgccgac tgggtcctgg atggacaacc ctggcgccacc 180

149/154

```

gtcagcctgg aggagccggt ctggaagcca gacatggggc tggtagccct ggaggctgaa 240
ggccaggagc tctgtcttga gctggagaag aaccatggcc tgatcacctt cagcaggaat 300
ggcagctatt atctgcgtcc ctggccaccc cggggctcca aggacttctc aaccacagag 360
at

```

```

<210> 354
<211> 518
<212> DNA
<213> Homo sapiens

```

```

<400> 354
gagggcagag agctcacagc tatgggctgg agggcccggg gagctcgggg gaccccggtg 60
ctgctgctgc tactactgct gctgctctgg ccagtgccag gcgcccgggg gcttcaagga 120
catatccctg ggcagccagt caccgcgcac tgggtcctgg atggacaacc ctggcgcacc 180
gtcagcctgg aggagccggt ctggaagcca gacatggggc tggtagccct ggaggctgaa 240
ggccaggagc tctgtcttga gctggagaag aaccacaggg tgctggcccc aggatacata 300
gaaacccact acggcccaga tgggcagcca gtggtgctgg cccccaacca cacggatcat 360
tgccactacc aagggcgagt aaggggcttc cccgactcct gggtagtcct ctgcacctgc 420
tctgggatga gtggcctgat caccctcagc aggaatgccg gctattatct gcgtccctgg 480
ccaccccggg gctccaagga cttctcaacc cacgagat

```

```

<210> 355
<211> 506
<212> DNA
<213> Homo sapiens

```

```

<400> 355
ctggccccag gatacataga aaccactac ggcccagatg ggcagccagt ggtgctggcc 60
cccaaccaca cggatcattg ccactaccaa gggcgagtaa ggggcttccc cgactcctgg 120
gtagtctctt gcacctgctc tgggatgagt ggctgatca cctcagcag gaatgccagc 180
tattatctgc gtccctggcc accccggggc tccaaggact tctcaacca cgagatcttt 240
cggatggagc agctgctcac ctggaaagga acctgtggcc acagggatcc tgggaacaaa 300
gcgggcatga ccagccttcc tggtagtccc cagagcaggg gcaggcgaaa agcgcgcagg 360
acccggaagt acctggaact gtacattgtg gcagaccaca cctgttctt gactcggcac 420
cgaaacttga accacaccaa acagcgtctc ctggaagtgc ccaactacgt ggaccagctt 480
ctcaggactc tggacattca ggtggc

```

```

<210> 356
<211> 503
<212> DNA
<213> Homo sapiens

```

```

<400> 356
cagtggctac tgctgggatg gcgcatgtcc cacgctggag cagcagtgcc agcagctctg 60
ggggcctggc tcccacccag ctcccagggc ctgtttccag gtggtgaact ctgcgggaga 120
tgctcatgga aactgcggcc aggacagcga gggccacttc ctgccctgtg caggagaggga 180
tgccctgtgt ggggaagctgc agtgccaggg tggaaagccc agcctgctcg caccgcacat 240
ggtgccagtg gactctaccg ttcacctaga tggccaggaa gtgacttgct ggggagcctt 300
ggcactcccc agtgcccagc tggacctgct tggcctgggc ctggtagagc caggcaccca 360
gtgtggacct agaattggtt gcaatagcaa ccataactgc cactgtgctc caggctgggc 420
tccacccttc tgtgacaagc caggcttttg tggcagcatg gacagtggcc ctgtgcaggc 480
tgaaaaccat gacaccttcc tgc

```

```

<210> 357

```

150/154

<211> 581  
 <212> DNA  
 <213> Homo sapiens

<400> 357  
 cagtggctac tgctgggatg gcgcatgtcc cacgctggag cagcagtgcc agcagctctg 60  
 ggggcctggc tcccacccag ctcccaggc ctgtttccag gtggtgaact ctgcgggaga 120  
 tgctcatgga aactgcggcc aggacagcga gggccacttc ctgccctgtg cagggaggga 180  
 tgccctgtgt gggaagctgc agtgccaggg tggaaagccc agcctgctcg caccgcacat 240  
 ggtgccagtg gactctaccg ttcacctaga tggccaggaa gtgacttgct ggggagcctt 300  
 ggcactcccc agtgcccagc tggacctgct tggcctgggc ctggttagagc caggcaccca 360  
 gtgtggacct agaattggtg gccagagcag gcgctgcagg aagaatgcct tccaggagct 420  
 tcagcgctgc ctgactgcct gccacagcca cggggtttgc aatagcaacc ataactgcca 480  
 ctgtgctcca ggctgggctc cacccttctg tgacaagcca ggctttggtg gcagcatgga 540  
 cagtggccct gtgcaggctg aaaaccatga caccttcctg c 581

<210> 358  
 <211> 380  
 <212> DNA  
 <213> Homo sapiens

<400> 358  
 cagtggctac tgctgggatg gcgcatgtcc cacgctggag cagcagtgcc agcagctctg 60  
 ggggcctgat ggccaggaag tgacttgctg gggagccttg gcactcccca gtgccagct 120  
 ggacctgctt ggccctgggccc tggtagagcc aggacccag tgtggacctt gaattggtgtg 180  
 ccagagcagg cgctgcagga agaattgcctt ccaggagctt cagcgctgcc tgactgcctg 240  
 ccacagccac ggggtttgca atagcaacca taactgccac tgtgctccag gctgggctcc 300  
 acccttctgt gacaagccag gctttggtgg cagcatggac agtggccctg tgcaggctga 360  
 aaacctatgac accttcctgc 380

<210> 359  
 <211> 324  
 <212> DNA  
 <213> Homo sapiens

<400> 359  
 ggcctggtgt tgctaccgac tcccaggagc ccatctgcag cgatgcagct ggggctgcag 60  
 aagggacctt gcgtgcagtg gcccacaaaga tggccacac agggaccacc ccctgggctg 120  
 cgttcacccc atggagttgg gcccacagc cactggacag ccctggcccc tggacctga 180  
 gaactctcat gagccagca gccacctga gaagcctctg ccagcagctt cgctgaccc 240  
 ccaagcagat caagtccaga tgccaagatc ctgcctctgg tgagaggtag ctctaaaaat 300  
 gaacagattt aaagacaggt ggcc 324

<210> 360  
 <211> 753  
 <212> DNA  
 <213> Homo sapiens

<400> 360  
 cagtggctac tgctgggatg gcgcatgtcc cacgctggag cagcagtgcc agcagctctg 60  
 ggggcctgat ggccaggaag tgacttgctg gggagccttg gcactcccca gtgccagct 120  
 ggacctgctt ggccctgggccc tggtagagcc aggacccag tgtggacctt gaattggtgtg 180  
 ccagagcagg cgctgcagga agaattgcctt ccaggagctt cagcgctgcc tgactgcctg 240  
 ccacagccac ggggtttgca atagcaacca taactgccac tgtgctccag gctgggctcc 300  
 acccttctgt gacaagccag gctttggtgg cagcatggac agtggccctg tgcaggctga 360



151/154

|            |            |            |             |             |             |     |
|------------|------------|------------|-------------|-------------|-------------|-----|
| aaaccatgac | accttcctgc | tggccatgct | cctcagcgtc  | ctgctgcctc  | tgctcccagg  | 420 |
| ggccggcctg | gcctggtggt | gctaccgact | cccaggagcc  | catctgcagc  | gatgcagctg  | 480 |
| gggctgcaga | agggaccctg | cgtgcagtgg | ccccaaagat  | ggccccacaca | gggaccaccc  | 540 |
| cctgggcggc | gttcacccca | tggagttggg | ccccacagcc  | actggacagc  | cctggccctt  | 600 |
| ggaccctgag | aactctcatg | agcccagcag | ccaccctgag  | aagcctctgc  | cagcagtctc  | 660 |
| gcctgacccc | caagcagatc | aagtccagat | gccaaagatcc | tgcctctggt  | gagaggttagc | 720 |
| tcctaaaatg | aacagattta | aagacaggtg | gcc         |             |             | 753 |

&lt;210&gt; 361

&lt;211&gt; 1154

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 361

|            |             |             |            |            |             |      |
|------------|-------------|-------------|------------|------------|-------------|------|
| cagtggctac | tgctgggatg  | gcgcatgtcc  | cacgctggag | cagcagtgcc | agcagctctg  | 60   |
| ggggcctggc | tcccacccag  | ctcccagggc  | ctgtttccag | gtggtgaact | ctgcgggaga  | 120  |
| tgctcatgga | aactgcggcc  | aggacagcga  | ggggcacttc | ctgccctgtg | cagggagggga | 180  |
| tgccctgtgt | gggaagctgc  | agtgccaggg  | tggaaagccc | agcctgctcg | caccgcacat  | 240  |
| ggtgccagtg | gactctaccg  | ttcacctaga  | tggccaggaa | gtgacttgte | ggggagcctt  | 300  |
| ggcactcccc | agtgccccagc | tggacctgct  | tggcctgggc | ctggtagagc | caggcaccca  | 360  |
| gtgtggacct | agaatggtgt  | gccagagcag  | gcgctgcagg | aagaatgcct | tccaggagct  | 420  |
| tcagcgctgc | ctgactgcct  | gccacagcca  | cggggtttgc | aatagcaacc | ataactgcca  | 480  |
| ctgtgctcca | ggctgggctc  | cacccttctg  | tgacaagcca | ggctttggtg | gcagcatgga  | 540  |
| cagtggccct | gtgcaggctg  | aaaaccatga  | caccttcctg | ctggccatgc | tcctcagcgt  | 600  |
| cctgctgcct | ctgctcccag  | gggcccggcct | ggcctggtgt | tgctaccgac | tcccaggagc  | 660  |
| ccatctgcag | cgatgcagct  | ggggctgcag  | aagggaccct | gcgtgcagtg | gccccaaaga  | 720  |
| tggccacac  | agggaccacc  | cctggggcgg  | cgttcacccc | atggagttag | gccccacagc  | 780  |
| cactggacag | ccctggcccc  | tggacctga   | gaactctcat | gagcccagca | gccaccctga  | 840  |
| gaagcctctg | ccagcagtct  | cgcctgaccc  | ccaagatcaa | gtccagatgc | caagatcctg  | 900  |
| cctctggtga | gaggtagctc  | ctaaaatgaa  | cagatttaaa | gacaggtggc | cactgacagc  | 960  |
| cactccagga | acttgaactg  | caggggcaga  | gccagtgaat | caccggacct | ccagcacctg  | 1020 |
| caggcagctt | ggaagtttct  | tccccgagtg  | gagcttcgac | ccaccactc  | caggaaccca  | 1080 |
| gagccacatt | agaagttcct  | gagggctgga  | gaacactgct | gggcacactc | tccagctcaa  | 1140 |
| taaaccatca | gtcc        |             |            |            |             | 1154 |

&lt;210&gt; 362

&lt;211&gt; 953

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 362

|            |            |            |            |             |            |     |
|------------|------------|------------|------------|-------------|------------|-----|
| cagtggctac | tgctgggatg | gcgcatgtcc | cacgctggag | cagcagtgcc  | agcagctctg | 60  |
| ggggcctgat | ggccaggaag | tgacttgctg | gggagccttg | gcactcccca  | gtgccagct  | 120 |
| ggacctgctt | ggcctgggcc | tggtagagcc | aggcaccag  | tgtggacctt  | gaatggtgtg | 180 |
| ccagagcagg | cgctgcagga | agaatgcctt | ccaggagctt | cagcgctgcc  | tgactgcctg | 240 |
| ccacagccac | ggggtttgca | atagcaacca | taactgccac | tgtgctccag  | gctgggctcc | 300 |
| acccttctgt | gacaagccag | gctttggtgg | cagcatggac | agtggccctg  | tgagggtga  | 360 |
| aaaccatgac | accttcctgc | tggccatgct | cctcagcgtc | ctgctgcctc  | tgctcccagg | 420 |
| ggccggcctg | gcctggtggt | gctaccgact | cccaggagcc | catctgcagc  | gatgcagctg | 480 |
| gggctgcaga | agggaccctg | cgtgcagtgg | ccccaaagat | ggccccacaca | gggaccaccc | 540 |
| cctgggcggc | gttcacccca | tggagttggg | ccccacagcc | actggacagc  | cctggccctt | 600 |
| ggaccctgag | aactctcatg | agcccagcag | ccaccctgag | aagcctctgc  | cagcagtctc | 660 |
| gcctgacccc | caagatcaag | tccagatgcc | aagatcctgc | ctctggtgag  | aggtagctcc | 720 |
| taaaatgaac | agatttaaa  | acaggtggcc | actgacagcc | actccaggaa  | cttgaactgc | 780 |
| aggggcagag | ccagtgaatc | accggacctc | cagcacctgc | aggcagcttg  | gaagtttctt | 840 |
| ccccgagtgg | agcttcgacc | caccactcc  | aggaaccag  | agccacatta  | gaagttcctg | 900 |

152/154

agggctggag aacactgctg ggcacactct ccagctcaat aaaccatcag tcc 953

&lt;210&gt; 363

&lt;211&gt; 812

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 363

Met Gly Trp Arg Pro Arg Arg Ala Arg Gly Thr Pro Leu Leu Leu Leu  
1 5 10 15

Leu Leu Leu Leu Leu Leu Trp Pro Val Pro Gly Ala Gly Val Leu Gln  
20 25 30

Gly His Ile Pro Gly Gln Pro Val Thr Pro His Trp Val Leu Asp Gly  
35 40 45

Gln Pro Trp Arg Thr Val Ser Leu Glu Glu Pro Val Ser Lys Pro Asp  
50 55 60

Met Gly Leu Val Ala Leu Glu Ala Glu Gly Gln Glu Leu Leu Leu Glu  
65 70 75 80

Leu Glu Lys Asn His Arg Leu Leu Ala Pro Gly Tyr Ile Glu Thr His  
85 90 95

Tyr Gly Pro Asp Gly Gln Pro Val Val Leu Ala Pro Asn His Thr Asp  
100 105 110

His Cys His Tyr Gln Gly Arg Val Arg Gly Phe Pro Asp Ser Trp Val  
115 120 125

Val Leu Cys Thr Cys Ser Gly Met Ser Gly Leu Ile Thr Leu Ser Arg  
130 135 140

Asn Ala Ser Tyr Tyr Leu Arg Pro Trp Pro Pro Arg Gly Ser Lys Asp  
145 150 155 160

Phe Ser Thr His Glu Ile Phe Arg Met Glu Gln Leu Leu Thr Trp Lys  
165 170 175

Gly Thr Cys Gly His Arg Asp Pro Gly Asn Lys Ala Gly Met Thr Ser  
180 185 190

Leu Pro Gly Gly Pro Gln Ser Arg Gly Arg Arg Glu Ala Arg Arg Thr  
195 200 205

Arg Lys Tyr Leu Glu Leu Tyr Ile Val Ala Asp His Thr Leu Phe Leu  
210 215 220

Thr Arg His Arg Asn Leu Asn His Thr Lys Gln Arg Leu Leu Glu Val  
225 230 235 240

Ala Asn Tyr Val Asp Gln Leu Leu Arg Thr Leu Asp Ile Gln Val Ala  
245 250 255

153/154

Leu Thr Gly Leu Glu Val Trp Thr Glu Arg Asp Arg Ser Arg Val Thr  
 260 265 270  
 Gln Asp Ala Asn Ala Thr Leu Trp Ala Phe Leu Gln Trp Arg Arg Gly  
 275 280 285  
 Leu Trp Ala Gln Arg Pro His Asp Ser Ala Gln Leu Leu Thr Gly Arg  
 290 295 300  
 Ala Phe Gln Gly Ala Thr Val Gly Leu Ala Pro Val Glu Gly Met Cys  
 305 310 315 320  
 Arg Ala Glu Ser Ser Gly Gly Val Ser Thr Asp His Ser Glu Leu Pro  
 325 330 335  
 Ile Gly Ala Ala Ala Thr Met Ala His Glu Ile Gly His Ser Leu Gly  
 340 345 350  
 Leu Ser His Asp Pro Asp Gly Cys Cys Val Glu Ala Ala Ala Glu Ser  
 355 360 365  
 Gly Gly Cys Val Met Ala Ala Ala Thr Gly His Pro Phe Pro Arg Val  
 370 375 380  
 Phe Ser Ala Cys Ser Arg Arg Gln Leu Arg Ala Phe Phe Arg Lys Gly  
 385 390 395 400  
 Gly Gly Ala Cys Leu Ser Asn Ala Pro Asp Pro Gly Leu Pro Val Pro  
 405 410 415  
 Pro Ala Leu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp  
 420 425 430  
 Cys Gly Pro Gly Gln Glu Cys Arg Asp Leu Cys Cys Phe Ala His Asn  
 435 440 445  
 Cys Ser Leu Arg Pro Gly Ala Gln Cys Ala His Gly Asp Cys Cys Val  
 450 455 460  
 Arg Cys Leu Leu Lys Pro Ala Gly Ala Leu Cys Arg Gln Ala Met Gly  
 465 470 475 480  
 Asp Cys Asp Leu Pro Glu Phe Cys Thr Gly Thr Ser Ser His Cys Pro  
 485 490 495  
 Pro Asp Val Tyr Leu Leu Asp Gly Ser Pro Cys Ala Arg Gly Ser Gly  
 500 505 510  
 Tyr Cys Trp Asp Gly Ala Cys Pro Thr Leu Glu Gln Gln Cys Gln Gln  
 515 520 525  
 Leu Trp Gly Pro Gly Ser His Pro Ala Pro Glu Ala Cys Phe Gln Val  
 530 535 540  
 Val Asn Ser Ala Gly Asp Ala His Gly Asn Cys Gly Gln Asp Ser Glu  
 545 550 555 560

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Gly | His | Phe | Leu | Pro | Cys | Ala | Gly | Arg | Asp | Ala | Leu | Cys | Gly | Lys | Leu |  |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |  |
| Gln | Cys | Gln | Gly | Gly | Lys | Pro | Ser | Leu | Leu | Ala | Pro | His | Met | Val | Pro |  |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |  |
| Val | Asp | Ser | Thr | Val | His | Leu | Asp | Gly | Gln | Glu | Val | Thr | Cys | Arg | Gly |  |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |  |
| Ala | Leu | Ala | Leu | Pro | Ser | Ala | Gln | Leu | Asp | Leu | Leu | Gly | Leu | Gly | Leu |  |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |  |
| Val | Glu | Pro | Gly | Thr | Gln | Cys | Gly | Pro | Arg | Met | Val | Cys | Gln | Ser | Arg |  |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |  |
| Arg | Cys | Arg | Lys | Asn | Ala | Phe | Gln | Glu | Leu | Gln | Arg | Cys | Leu | Thr | Ala |  |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |  |
| Cys | His | Ser | His | Gly | Val | Cys | Asn | Ser | Asn | His | Asn | Cys | His | Cys | Ala |  |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |  |
| Pro | Gly | Trp | Ala | Pro | Pro | Phe | Cys | Asp | Lys | Pro | Gly | Phe | Gly | Gly | Ser |  |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |  |
| Met | Asp | Ser | Gly | Pro | Val | Gln | Ala | Glu | Asn | His | Asp | Thr | Phe | Leu | Leu |  |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |  |
| Ala | Met | Leu | Leu | Ser | Val | Leu | Leu | Pro | Leu | Leu | Pro | Gly | Ala | Gly | Leu |  |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |  |
| Ala | Trp | Cys | Cys | Tyr | Arg | Leu | Pro | Gly | Ala | His | Leu | Gln | Arg | Cys | Ser |  |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |  |
| Trp | Gly | Cys | Arg | Arg | Asp | Pro | Ala | Cys | Ser | Gly | Pro | Lys | Asp | Gly | Pro |  |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |  |
| His | Arg | Asp | His | Pro | Leu | Gly | Gly | Val | His | Pro | Met | Glu | Leu | Gly | Pro |  |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |  |
| Thr | Ala | Thr | Gly | Gln | Pro | Trp | Pro | Leu | Asp | Pro | Glu | Asn | Ser | His | Glu |  |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |  |
| Pro | Ser | Ser | His | Pro | Glu | Lys | Pro | Leu | Pro | Ala | Val | Ser | Pro | Asp | Pro |  |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |  |
| Gln | Asp | Gln | Val | Gln | Met | Pro | Arg | Ser | Cys | Leu | Trp |     |     |     |     |  |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     |     |     |  |